

现代经济学研究丛书

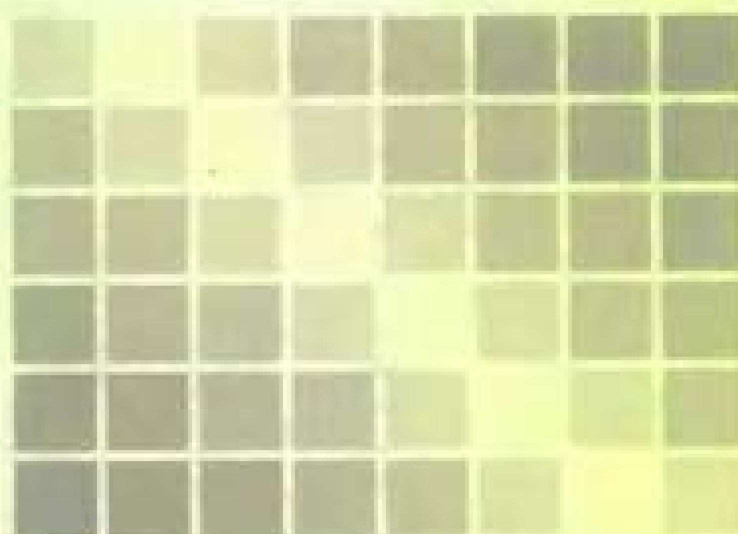
丛书主编　费方域

博弈论

施锡铨　著

XIANDAI JINGJIXUE YANJIU CONGSHU

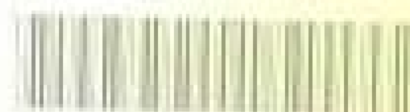
上海财经大学出版社





VISIT…

LANZAROTE Caliente.com



财经大学 2 070 9605 1

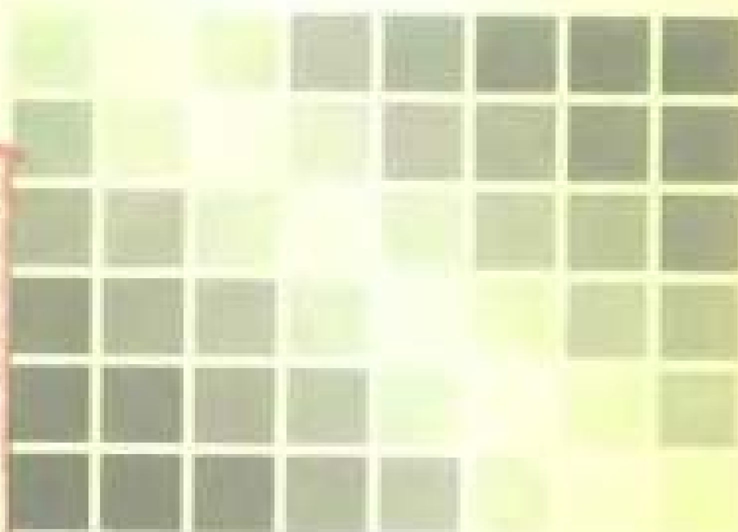
博弈论

施锡铨 著

XIANDAI JINGJIXUE YANJIU CONGSHU

上海财经大学出版社





图书在版编目(CIP)数据

博弈论/施锡铨著．—上海：上海财经大学出版社，2000.2
(现代经济学研究丛书)
ISBN 7-81049-398-1/F·334

Ⅰ.博… Ⅱ.施… Ⅲ.对策论 Ⅳ.0225

中国版本图书馆 CIP 数据核字(2000)第 11089 号

丛书主编 费方域
责任编辑 江 玉
封面设计 未 名

BO YI LUN
博弈论
施锡铨 著

上海财经大学出版社出版发行
(上海市中山北一路 369 号 邮编 200083)
全国新华书店经销
上海第二教育学院印刷厂印刷
上海浦江装订厂装订
2000 年 2 月第 1 版 2000 年 2 月第 1 次印刷

850mm×1168mm 1/32 14.625 印张 366 千字
印数 1—3000 定价：32.00 元

总 序

这是一套侧重介绍和研究现代微观经济学的丛书。它的内容将尽可能地涉及有关的基础理论知识和应用理论知识已经和正在不断地丰富、拓展和推进着的脉络、框架、边界，与亟待吸取、回答和重视的经验、问题和政策。具体来说，或多或少地，大致包括：博弈理论，经济建模理论与方法，一般均衡理论，竞争理论与政策，寡头理论与致策，信息、激励与合同理论，管制、放松管制与再管制的理论和方法，最优税收理论，交易成本理论，比较制度与转型经济的理论和分析，外部效应理论，公共产品理论，企业理论，银行金融理论和公司财务理论等。

追溯这些理论的产生和演进，不难发现：它们的孕育，主要集中在即将揖别的这个千年的后一二百年里，并且随着时间每推移三五十年而呈现加速积累；它们的策源地，主要集中在西欧，特别是北美。这并不奇怪。因为正是在此期间，这些地方率先进行了工业革命和信息革命，率先建立了市场机制和相应的法律制度，率先实现了经济起飞和持续增长，率先获得了繁荣和富裕。而经济学作为理论，它的形成有待于经济实践的展开和需求；经济学作为知识，它的创造需要良好的环境和高昂的成本；经济学作为方法，它的运用依赖于其他科学的发展和完善。这些地方在经济、科技、文化等方面的先动优势，造成了它们在知识从而在经济学方面的先动优势，造成了它们与其他地方在经济、科技、文化、知识从而在经济学方面的差距和不公平分配。而这两种优势的互动，又在不断扩

我本从事统计学工作，对于博弈论，只是略知皮毛。但是我给学生布置的博士学位论文却使自己“老革命遇到新问题”，诸如可持续发展经济、企业的兼并和重组、可转换公司债券、粮食流通体制改革等等，不仅需要统计学的知识，而且也涉及到博弈理论。学生对这方面的要求迫使我下决心“销声匿迹”于统计学界，“闭关”两年多，重新“修炼”。在这个过程中，边学边教。有时自己感到不知所云，且学生听得目瞪口呆，就知道“道行”不够，只得面壁重思，直到学生能够理解。渐渐地，他们眉开眼笑的机会越来越多，我觉得自己的体会越来越深。忽然一个想法涌上心头：为什么不把这些心得总结一下呢？于是就有了这本书。

一次，有个学生拿了一篇有关农业生产的博弈分析论文给我看。这是一篇很成问题的论文。学生想不通为什么一本在数量经济方面颇有知名度的杂志会刊登这样错误的文章。我想，真理在被人们认识的过程中，很有可能发生这类事情。其实，那篇文章主要错在基本概念。这件事情对本书的写作增加了无形的压力，生怕一着不慎而误人子弟。有些内容不得不从本书的原稿中删去，因为有些数学内容如何演算成为经济类读者容易理解的东西，正是作者近来苦思冥想尚无把握的事情。如果本书有再版的可能，希望经过探索，学有所成，能将这部分内容补充进去。

本书写作过程中遇到经费上的困难，一度使我十分沮丧。感谢谈敏教授、费方域教授、丛树海教授等对本项工作的一贯关心与鼓励，使我重新鼓起勇气正视现实。有段时候学起和尚“到处化缘，以修佛身”的手段，敲起木鱼，以感动各方善男信女。令人感动的是，上海财经大学“211 工程”了解到该情况后，立即对《博弈论》的写作给予极大支持和资助，并追加立项，得以使《博弈论》顺利完成，佛像终于着了金装。根据积累的资料和新近从互联网上学到的东西，我觉得也许还可以写点“博弈论解题技巧”、“机制设计原理”、“近代博弈论”等，尤其是解题之类的内容对于初学者掌握这门知

全，状况并不确定，信息并不对称，决策并不独立，合同并不完全，交易并不免费，资源并不都由价格配置的时候；当人们思索为什么一部分经济活动要由政府、企业、银行等金融机构来组织，为什么法律、习惯、机构等制度和预期、风险、管理、机制、金融工具等因素会对人的行为和相互关系从而对经济绩效发生影响的时候；更一般地说，当人们逐渐发现原有理论的假设或结论不能被实际情况和观察数据证实（证伪），或者，原有理论从观察中获得了新的概念、新的方法，或者，数学工具的飞跃发展可以为理论提供更贴切、更明白的模型，可以为复杂难懂的问题提供更为严谨（合理）的解的时候，对新古典传统的怀疑、责难、修正和批评就不断地发生了。于是，在确定性情况下的选择和最优理论之外，创立了不确定性情况下的选择和最优理论。在瓦尔拉斯均衡和帕累托效率概念之外，形成了纳什均衡以及有关纳什均衡的一系列精炼的均衡概念。在完全信息完全合同的分析之外，发展了不完全信息不完全合同情况下的委托代理、机制设计和产权理论。在微积分等古典数学工具的运用之外，又大量应用了概率论与数理统计、时间序列分析、统计决策与贝叶斯理论等近代数学工具，以及企业理论、公共产品和公共选择理论，等等。它们或改写，或新创，从而积累了本丛书想要反映的上述众多内容。有意思的是，与 19 世纪末不同，这次，捍卫者没有采取终结真理的态度，创新者也没有采取全盘否定过去的态度。结果，双方相互理解、促进、吸取和包容，仔细地界定出每个原理、每项方法的创始人，以及它们的作用条件与范围。因此，我们希望本丛书的一个成功之处，是对此作出精确的描述。

1987 年以来，中国的融入国际市场一体化的改革，已经迈出了不可逆转的一步，并取得了举世瞩目的成绩。现在的问题，最根本的，一个是要回答市场社会主义道路是否能够和怎样才能走得通，另一个是要回答经济的可持续增长是否能够和怎样才能办得到。这两个问题，涉及企业、银行、财税改革，资产重组和资本市场

发育，经济发展战略和经济结构调整，科技和工业创新，通货膨胀和通货紧缩防治等诸多方面，复杂而又紧迫。寻找这些问题的答案，用得着上面提到的所有经济学知识。摸索路子的过程，就是我们缩小经济学方面与国际先进水平差距的过程，也是我们对属于全人类的经济科学作出自己贡献的过程。我们希望，本丛书最大的光荣，在于它能多多少少地留下一些中国人不畏艰险、不辞辛劳地攀登经济和经济学现代化巅峰的勇气、智慧和足迹。

新世纪就在前头，让我们用耕耘迎接她的到来。

费方域

1999 年 12 月

前 言

博弈论距离中国人的生活越来越近了。自从张维迎教授的“博弈论与信息经济学”问世以来，在中国的经济学界和管理学界，有关博弈论的应用，非但是听到了雷声，而且有时还看到了雨点。

博弈论可以广泛地应用于包括政治、经济以及军事等各个领域。1999 年 4、5 月间，我曾就北约轰炸南斯拉夫建立了一个粗糙的博弈模型。发现北约寻求的均衡解是：北约一轰炸，南斯拉夫就投降。但是它忽略了非均衡途径的情况：轰炸不能迫使南斯拉夫投降。读者将从本书中获知，忽略非均衡途径的策略不是完美的。因此，当时可以预测北约有可能陷入一个窘境：它一方面继续轰炸，一方面谋求第三者出面调停。造成这种局面的原因多种多样，交战双方对盈利函数估价的不同是其中的原因之一：北约认为轰炸将使南斯拉夫造成经济上的重大损失，而南斯拉夫认为国家的主权与领土完整高于一切。由于资料的缺乏，我们无法从简单的模型中预测后来的结局，但是博弈论的功效在这里是非常明显的。

严格地说，本书只是作者对书末所列参考文献的学习笔记，正因为有些心得体会，因此细心的读者也许会发现书中的某些证明与解题同原版书有所不同。事实上，我对发生在原版书中的错误论证作了(自以为是的)改正工作。

博弈论专著理应由博弈论专家来写。我认为他们应当具有哲学家的头脑、数学家的技巧和经济学家(或其他领域专家)的丰富知识。至于我，正如平时开玩笑所说，仅仅是“上完了课就回家”。

GDD 56/21

大这种差距和加深这种不公平。其他后起地方要想改变这种格局，既靠不得乞求，也犯不着愤懑和嫉妒。最理性的办法是学习和赶超——在引进发达国家的资金和先进技术、先进管理经验的同时，主动地大量收集和获取、吸收和消化先进的经济学知识，并进一步发展自己的理论。知识是人类共有的财富。已有的知识，可以通过各种途径得到交流和转让。新科技尤其是通信技术的发展，在带来新知识爆炸式增长的同时，也带来了知识迅速传播的便利和低成本获取的机遇。市场化（包括机构、政策、规则等）的深入，教育的普及和提高，在带来激烈竞争的同时，也带来了选择和接受知识的激励、能力、机制和效率。因此，正如我们完全应该也完全能够向一切先进的技术和先进的知识学习一样，我们也完全应该和完全能够向一切先进的经济学学习。按照博弈理论，有时候，后动也有优势，关键在于找准路子。我们编著这套丛书的一个重要的目的和依据，就是想为此尽一份绵薄之力。

回眸百年，我们都还记得，1970 年，首届诺贝尔经济学奖获得者拉格纳·弗里希的公开演讲，是从对 19 世纪经济学发展的简短考察开始的。他指出，那时，斯图阿特·穆勒认为，他对古典经济学的总结，已经使之成为一个有机的、逻辑和外形都臻于完全的整体，因此，价格和价值的一般理论都不再能增添什么了。但后来的发展否定了这个独断。一个突破来自边际主义革命，新古典经济学用主观因素观点补充了古典经济学的生产成本观点；另一个突破来自计量经济学，粗糙、简单、定性的东西，被经济理论、数学和统计三位一体的定量内容所替代。现在，当我们瞻望 20 世纪经济学的历程的时候，我们又一次领悟到知识和科学的发展永无止境的道理。这期间，一方面，经由萨缪尔逊和阿罗等人的探索，新古典经济学的核心体系更加严密、精巧、一以贯之，从而对斯密“看不见的手”的经验直觉赋予系统的形态和科学的性质，为市场经济的运作框架和秩序奠定了理论基石。另一方面，当人们发现市场并不完

识非常重要。但是我很担心是否又要抄起木鱼来。唉，愁煞人！

我这所谓的心得体会自然难免有所差错，恳盼读者指正，以帮助作者更上一层楼，也使其他读者免受误导之苦。

施锡铨

1999 年夏于上海财经大学

目　录

总　序/1

前　言/1

第一章　引论

§1.1　长街上的超市/1
§1.2　共同投资问题/4
§1.3　什么是博弈论/5
§1.4　博弈的分类/8

第一部分　完全信息静态博弈

第二章　策略型博弈与 Nash 均衡

§2.1　两人零和博弈——猜谜/11
§2.2　混合策略/13
§2.3　累次严优/18
§2.4　累次严优的应用实例/26
§2.5　Nash 均衡/30
§2.6　多重 Nash 均衡/42
§2.7　相关均衡/53

第三章 Nash 均衡存在性定理

§3.1 Cournot 竞争的 Nash 均衡可视作市场调整的结果/58
§3.2 Brouwer 不动点定理/60
§3.3 Sperner 引理/68
§3.4 Kakutani 不动点定理/71
§3.5 Nash 均衡存在性/77
§3.6 连续盈利无限博弈中的 Nash 均衡存在性/81

第二部分 完全信息动态博弈

第四章 展开型博弈

§4.1 定义与博弈树/89
§4.2 展开型博弈的策略与均衡/93
§4.3 Stackelberg 博弈与后退归纳法/110
§4.4 子博弈和子博弈完美均衡/115
§4.5 关于后退归纳法与子博弈完美的评论/118

第五章 多阶段博弈子博弈完美的应用

§5.1 可观察行动多阶段博弈/125
§5.2 有限范围博弈的一阶段偏离准则/127
§5.3 具有静态均衡的重复博弈/132
§5.4 两阶段博弈的若干经济应用/137
§5.5 消耗战与占先博弈/147
§5.6 开环和闭环均衡/156
§5.7 有限与无限状态均衡/160

第六章　讨价还价模型
§6.1　不存在耐心问题的讨价还价/163
§6.2　无耐心的讨价还价/166
§6.3　讨价还价的一般模型/169
§6.4　序贯两人讨价还价的实验迹象/174

第七章　宏观经济模型
§7.1　宏观经济模型/179
§7.2　宏观动态博弈的均衡解/184

第八章　重复博弈
§8.1　有限重复博弈/186
§8.2　无限重复博弈与无名氏定理/194
§8.3　无限重复博弈的若干例子/220
§8.4　对手变化的重复博弈/234
§8.5　Pareto 完美均衡与重新谈判检验/245
§8.6　公共信息不完美时的重复博弈/254

第三部分　不确定结局的博弈问题

第九章　道德风险问题
§9.1　道德风险与不完全保险/259
§9.2　道德风险与非自愿失业/271

第四部分　不完全信息静态博弈

第十章　Bayes 博弈与 Bayes 均衡

§10.1　非对称信息下的 Cournot 竞争/282
§10.2　Harsanyi 转换/285
§10.3　Bayes 均衡/289
§10.4　Bayes 均衡的若干例子/291
§10.5　混合策略的再解释/305

第十一章　机制设计

§11.1　机制设计的若干例子/316
§11.2　显示准则与机制设计/327

第五部分　不完全信息动态博弈

第十二章　完美 Bayes 均衡

§12.1　完美 Bayes 均衡定义/336
§12.2　具可观察行动与不完全信息的多阶段博弈/345
§12.3　序贯均衡/353
§12.4　颤抖手完美均衡/365
§12.5　适度均衡/375

第十三章　信号博弈

§13.1　信号博弈中的完美 Bayes 均衡/379
§13.2　劳务市场信号博弈/387
§13.3　合作投资与资本结构/400

§13.4　廉价交谈博弈/403

第十四章　完美 Bayes 均衡应用

§14.1　有限重复囚徒窘境的信誉效应/415

§14.2　非对称信息下的序贯讨价还价/424

第十五章　完美 Bayes 均衡的精炼

§15.1　剔除严劣策略/439

§15.2　"啤酒与蛋奶火腿蛋糕"信号博弈/445

§15.3　直觉准则/449

参考文献/452

第一章　引　论

§1.1　长街上的超市

“月亮为什么跟着人走?”“天上的云为什么飘在空中不会掉下来?”这是小孩常常会向大人提出的问题,表现出他们强烈的求知欲望。为此,出版社组织了大批专家撰写了大量通俗易懂的回答文章,这就是众所周知的少儿读物《十万个为什么》,它曾经风靡全国,一版再版。这套书基本上涉及的是自然科学领域中的问题,其实在社会科学领域,尤其是经济管理方面,存在着大量问题值得身处市场经济中的人们去思考、去探求。

例如,某报就提出过一个有关超市在商业街的布局问题,并对此进行了讨论。细心的人们常常发现在不少较长的街上,华联、联华、农工商和其他一些超市似乎“喜欢”拥挤在一起。有人指责这属于“资源浪费”现象,因为他们设想在较长的街的两旁较均匀地分布(或几乎等距离地分布)若干个超市,这对周围地区的居民无疑构成了极大的方便。为了回答这样的问题,我们不妨来观察一个更为简单的现象。

据说西方发达国家的不少男男女女有日光浴的爱好,因为它有利于身体健康。现在设想较长的海滩上比较均匀地散布着许多日光浴者。这里,之所以采用“比较均匀地散布”这样一个假设,是考虑到人过分地拥挤在一起将可能感到难受且达不到日光浴的效

果。太阳的照射使人们需要补充水分。假如有 A 与 B 两个小贩来到海滩，以同样的价格、相同的质量向日光浴者提供同一品牌的矿泉水（或啤酒）。在直线状的海滩上他们应当如何合理地安置自己的摊位呢？若将海滩长度标准化为 1，那么图 1.1 中的[0,1]线段则表示海滩，“ * ”号则代表那些日光浴者。

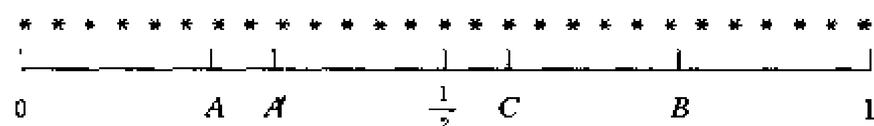


图1.1

再作一个合乎逻辑的假定：通常情况下，日光浴者总是到距自己最近的摊位购买矿泉水。根据这个原则，不少人将会赞成如下的摊位安排：如果以 0 为起点，A 在 1/4 处，B 在 3/4 处。这样既方便了众多日光浴者，而且 A、B 两人各占约一半顾客的生意，可谓公平合理、皆大欢喜。

然而，生意人都有自己的“理性”，只要手段合法，总是希望自己的生意尽可能地红火，至于他人生意的好坏则不管自己的事。正是出于这种理性，小贩 A 自然地产生如下想法：如果我将自己的摊位往 B 那儿挪动一下至 A' 位置（仍见图 1.1），那么从 0 至 A' 范围内的人显然是我的顾客，而 A' 与 B 之间的中点将从原来的 1/2 处移至 1/2 右边的 C 处，从 A' 至 C 范围内的人也将成为我的顾客，这样无疑地我从 B 那儿“夺”走了一部分生意。毋庸置疑，这是个好主意！然而，B 也是一个“理性”的商人，这种简单地扩大顾客数目的办法他不会想不到，因此 B 自然也会将自己的摊位往 A 的方向挪动。不难想象，双方“斗智”的结果将使 A、B 两个小贩的摊位都移到海滩中点 1/2 处相互为邻，并相安无事地做起他们的矿泉水买卖。

这个现象实际上是两个理性人采用理性行为的自然结果。回到长街上的超市现象，如果地段的繁华等其他原因可以认为相同

的话，那么，只要条件许可，超市的几乎相依为邻现象完全可以看作公正的市场竞争的合理结果。

在社会经济领域内，有不少现象与上述小贩的摊位安置、长街上的超市位置有着相似之处，从某种意义上也可以用上述逻辑进行分析与阐述。

(1)同一城市的两家航空公司开辟飞往同一目的地的航班，常出现他们各自的起飞时刻被安排在几乎同一时间的现象。

(2)人们对电视节目的喜爱存在着一定的档次差异，因此电视台对节目的编排将直接影响到收视率。设想如果将高雅艺术节目与较低档趣味的节目比作海滩的两端，那么观赏电视节目的观众就相当于散布在海滩上的日光浴者。可以想象"各电视台为争取尽可能多的观众"类似于"卖饮料的小贩在努力争取尽可能多的顾客"。因此不少电视台常将黄金播放时段的文艺节目定位于中等趣味以提高自己的收视率。

此外，各电视台中一些内容虽然不同但情调却差不多的娱乐节目，常在播放时间上撞车。也可以利用海滩占位模型作出相应的解释。

(3)海滩占位问题在政治学中也可以找到类似的案例。倘若将人们的极左与极右两种政治倾向视作海滩的两端，那么一个国家的选民相当于分布在海滩上的日光浴者(当然，分布未必均匀，通常持中间态度的人较多一些)。将竞选纲领定位于偏中相当于摊位选择在海滩中间。四年一次的美国总统竞选中，无论是民主党，还是共和党，其总统竞选人花费很大功夫推出自己的竞选纲领，目的旨在争取尽可能多的中间立场的选民。

上述种种，小到超市的布局，大到美国总统的竞选，我们将他们归结为海滩占位模型。这相当于两个(或两个以上的)小贩在玩一场抢占有利摊位的"游戏"，目的非常明确，他们希望极大化自己的利益。为此目的，他们必须考虑并预测游戏的对手可能怎样做，

从而选定对自己有利的决策。在社会经济领域内，许多事情可以归结为一场“游戏”，当然他们并不都是海滩占位模型。下面给出了另一类“游戏”。

§1.2 共同投资问题

设想有两位投资者，共同投资一个较大的项目，他们可以获得较大的回报。如果他们俩中至少有一个抽出资金用于一个小项目投资，他肯定可以获得相应回报，但比投资较大项目时获益要小得多。然而他的这一做法将导致较大项目陷入困境，从而使另一位投资者蒙受损失。是冒一定风险坚持投资较大项目以获得较大回报，还是抽回资金投资小项目至少有个“旱涝保收”，这就需要投资者作出决策并实施行动。

对于这样一个问题，我们也可以先考虑一个类似的狩猎“游戏”：两个猎人围住了一头鹿，他们各卡住鹿可能逃跑的两个关口中的一个。只要他们齐心协力，鹿就会成为他们的猎物。如果此时周围跑过一群兔子，两位猎人中的任何一个只要去抓兔子一定会获得成功，他会抓到一只小兔，但鹿却从他所把守的关口逃跑。现在他们必须同时作出决定：是猎鹿还是抓兔子。这是一个简单的游戏，游戏的结局不外乎是：二人合力猎鹿并平分一头鹿；二人都各自去抓兔子并各人有一只兔子进账；一人去抓兔而另一人坚持猎鹿，那么抓兔者猎获一只兔子而另一位则两手空空。连小学生也能知道，第三个结局是三种结局中最差的一种。通常，半头鹿比兔子值钱，因此第一种结局对两个猎人来说也许是最好的，但是，谁也很难否定两人都去抓兔子是很不错的结局，因为毕竟每个人都有所获。那么究竟以何种结局作为这场狩猎游戏的预测比较合理呢？事实上，如果没有更多的信息——诸如猎人的习性，对猎物价值的评估，或者对自己获益的期望等等——合理地预测结局将是十分

困难的。

同样，在二人共同投资问题上，两位投资者都希望了解投资较大或较小项目中各自可能的获益、对手的习惯以及对投资获益的期望等等，在了解尽可能多的情况下，他们才有可能同时作出自己的最佳决策。

用游戏模式来刻画经济管理现象在商业活动中尤为普遍。因为商业本身就是“游戏”，它是世界上人们所知道的“如何去玩”的最大游戏。例如，在两个公司关于市场占有的竞争就正如下棋一样，有着特定的规则，在两个对手的行为以及最终结局之间存在着一定的关系。但是不存在某一方完全控制最终结局的情况。本书目的在于引进和研究经济管理领域中的“游戏”模型，即所谓博弈论。

§1.3 什么是博弈论

前面两节中引入的游戏模型存在着一些共同的特点：每个游戏常有两个以上的参加者，他们在游戏中都有着自己的切身利益，今后我们称他们为局中人，常用记号 $i=1,2,\cdots$ 来表示。每个局中人都有着自己的可行行动集供自己选择，这种选择毫无疑问地会影响到其他局中人的切身利益。游戏中的各个局中人理性地采取或选择自己的策略行为，使得在这种相互制约、相互影响的依存关系中，尽可能地提高自己的利益所得，这正是游戏理论的关键所在。就像下棋游戏中的各方使尽浑身解数使自己尽可能赢或至少不输一样。将英文“The Game Theory”翻译成“博弈论”，其原因盖出于此。在本书的以后部分，我们经常将“一场游戏”、“游戏的局中人”等中的“游戏”两字说成“博弈”，以使叙述起来方便，希望不致引起读者的混淆。

现在，我们试图用一句话概括什么是博弈论：博弈论就是关于

包含相互依存情况中理性行为的研究。

所谓相互依存，通常是指博弈中的任何一个局中人受到其他局中人的行为的影响，反过来，他的行为也影响到其他局中人。由于这种相互依存性，游戏或博弈的结果依赖于每一个局中人的决策，没有一个人能完全地控制所要发生的事情，也没有一个局中人处于孤独的状态。相互依存常使博弈中的局中人之间产生竞争，譬如两个人分蛋糕，每个局中人都希望自己的那块可以分得大一些。然而，竞争仅仅是博弈论中相互依存的一个方面，应该指出，通常地博弈并非纯粹是局中人之间的竞争，相互依存的另一个方而是局中人可以有某些共同的兴趣或利益所在。仍以分蛋糕为例，作为局中人策略行动的结果，蛋糕的大小可以增加或者减少。局中人的共同兴趣在于增加蛋糕的总量，他们互相"倾轧"之处在于如何分配。从博弈论研究的角度，增大蛋糕应是博弈的第一步，而分配蛋糕则是博弈的第二步。

酝酿已久并正在逐步实施中的宝山钢铁公司与上海钢铁集团之间的"强强联合"存在着经济博弈的内容，这里而就有做大蛋糕与分配蛋糕的文章。据说，由于客观上存在着宝钢在效益、福利、待遇等各方而优于上钢的事实，许多宝钢职工对联合待不积极态度，担心自己的利益会因此而受到影响，而不少上钢职工则持欢迎态度，指望通过强强联合改善自己的生活待遇。其实，宝钢与上钢的强强联合，首先就是应考虑到如何发挥双方各自的优势，使联合后的企业能跃上一个新的台阶。或者说就是要先做大蛋糕，完成这一步以后，对于更大蛋糕的分配将会使国家与双方的职工均获得较为满意的份额。

在博弈论中还需要对一个词"理性行为"作一些说明。博弈论中的所谓理性，一般不是指道德标准。从参加博弈的局中人的眼光来看，他们试图去实施自己认为可能最好的行为，尽管这样的行为有可能损害了其他局中人。例如海滩占位问题中 A 将自己的摊位

往 B 的摊位挪动，这一举动损害了 B 的利益，因为它事实上造成 A 从 B 那儿夺走部分顾客，但是，从 A 的观点，这一举措提高了自己的收入，博弈论研究（和通常的经济研究）认为这是理性行为。因此“理性行为”似乎有点“利己，不管它是否损人”。从博弈分析人员的“旁观”眼光，一般不会去擅自判断局中人的动机究竟如何。或许人们认为局中人的目的有点近乎于疯狂，但是如果局中人是有意图地、系统地继续行动的话，我们认为他是理性的。理性的局中人至少不会持续地犯相同的错误！但是我们还必须强调，博弈问题中的理性一定要有合法的前提，对于那些非法的交易或买卖，尽管当事人在交易的过程中也是尽可能地极大化自己的利益而采取自己认为的最佳行动，但由于交易的非法性而使得博弈论不去涉及这样的事例。最有代表性的非法交易当数买卖毒品，讨价还价的结果是卖者赚到“昧心钱”，而使吸毒者得到“心满意足”的“享受”，博弈理论认为这类非法行为绝非理性，故而不加以讨论。

由于局中人的相互依存性，博弈中一个理性的决策必定建立在预测其他局中人的反应之上。一个局中人将自己置身于其他局中人的位置并为他着想从而预测其他局中人将选择的行动，在这个基础上该局中人决定自己最理想的行动，这就是博弈论方法的本质与精髓。

博弈论中每一个局中人作出理性决策的重要依据之一是他的可能盈利有多少，这就是一个局中人需要认真计算的盈利函数（payoff function）。对于每一个局中人，如果他们在可供自己选择的策略空间中任取一策略作为自己的行动，既不会给自己带来盈利，又不会使他们必须付出，这种失去了激励机制的游戏本身也就失去了“博”的意义，在社会经济领域中尤其不太可能出现这类现象。盈利函数的结构与取值无疑将会影响到局中人的行为，因而也影响到博弈的最终结局。由此，盈利函数的确定在博弈论研究中是件非常重要的事情。从对博弈的不同角度考虑，从局中人不同的观

点出发可以有形形色色的盈利函数。譬如夫妻俩一同去买菜，从时间上来看是一种浪费，从经济角度可能“支付”过大。但是如果从增强双方感情来定义盈利或效用函数，那么他们会作出“一起去买菜”的决策。

§1.4 博弈的分类

从博弈论的基本刻画可以看出，无论什么样的博弈，总是存在着如下三个要素：

(1)局中人，以 $i=1,2,\cdots$表示。

(2)每个局中人一般有若干个策略(strategies)可供选择，它们构成了该局中人的纯策略空间。局中人 i 的纯策略空间用 S_i表示，倘若 S_i 由 k_i 个纯策略构成，则有 $S_i=(s_{i1},s_{i2},\cdots,s_{ik_i})$。

纯策略空间有时也可以是连续的，比如在 AB 线段的海滩上摊位的选择就可认为几乎是连续的。

(3)每个局中人的盈利函数。我们记局中人 i 的盈利函数为 $u_i(s)$，其中 $s=(s_1,\cdots,s_r)$，而 s_j 表示局中人 j 所取策略，s 表示 r 个局中人的策略向量。显然，盈利函数 $u_i(s)$与 s 有密切关系。它是每个局中人真正关心的东西。

策略空间、盈利函数以及局中人的与博弈有关的特征等知识构成博弈的信息，从信息的角度，博弈可以分为完全信息与不完全信息两类。所谓完全信息是指每一个局中人对于自己以及其他局中人的策略空间、盈利函数等知识有完全的了解，否则，博弈就是不完全信息的。

博弈的分类还可以从局中人行动的先后次序着手，如果局中人同时选择行动，则称博弈为静态的。要求“同时”并不一定等于规定在同一时刻大家一起行动。通常在时间上虽有行动的先后，但局中人彼此不知道其他人在采取什么具体行动(直到博弈结束时)，

其效果仍等价于他们在同时行动，此时我们仍称它是静态博弈。倘若局中人的行动有先后顺序，后行动者可以观察到前行动者的行动，并在这基础上采取自己最有利的策略，博弈就转为动态形式。

上面两种划分类型中各自存在两种情况，如果两两交叉，就可以得到完全信息静态博弈、完全信息动态博弈、不完全信息静态博弈和不完全信息动态博弈四种情况。无可非议，不完全信息要比完全信息复杂一些，而静态的考虑远比动态简单得多。因此本书将按上述顺序，从较简单情况入门逐一介绍。另外考虑到经济博弈的特点，我们在第九章中稍微讨论了结局为不确定的博弈。

第一部分　完全信息静态博弈

第二章　策略型博弈与 Nash 均衡

§2.1　两人零和博弈——猜谜

我们首先介绍最简单的猜谜游戏，从而逐一引进完全信息静态博弈的基本概念。

这是儿童常玩的一种游戏。有甲、乙两人，当说“开始”时，乙要么出示一个指头、要么出示两个指头，与此同时，甲也在这两种可能中作出选择出示一个或两个指头。游戏规则：如果两人各出示的指头数完全相同，则规定甲赢，乙必须付给甲 1 元；倘若他们所出示的指头数不匹配，则算乙赢，甲必须付给乙 1 元。博弈共有四个可能结局：(甲，乙)＝(1，1)，(1，2)，(2，1)和(2，2)。

猜谜博弈中共有两个局中人甲与乙，分别以 1 与 2 表示之。

每个局中人的纯策略空间 $S_1=S_2=(1,2)$。

对于局中人 1 与局中人 2 的每一种策略组合(即博弈的结局)，他们各有自己的盈利：

$$u_1(1,1)=1,u_1(1,2)=-1,u_1(2,1)=-1,u_1(2,2)=1$$

$u_2(1,1)=-1, u_2(1,2)=1, u_2(2,1)=1, u_2(2,2)=-1$

如果我们把以上盈利用一个矩阵表示，则更加一目了然（见图2.1）。

局中人 2

		1	2
局中人 1	1	1,−1	−1,1
	2	−1,1	1,−1

图 2.1

上述矩阵称为盈利矩阵，矩阵外所标的 1 与 2，分别表示局中人 1 与 2 所采用的策略与行动。矩阵的每一格相应于局中人的策略组合，格中的数字即为局中人在该结局中应得的盈利。为方便起见，习惯上规定左边的数表示矩阵左侧局中人 1 的盈利，右边那个数表示矩阵上方的局中人 2 的盈利。

可以设想猜谜博弈中的局中人可以是两个队、两家公司或者两个国家。类似于猜谜博弈的模型至少有一个明显的特点：在博弈的每一个结局中，两个局中人的盈利之和恰好等于零。凡具有这一特性的两人博弈统称为两人零和博弈。一般博弈论教科书中将两人零和博弈的定义稍稍扩大一些，如果两人博弈中每个结局里局中人盈利之和等于常数，也称该博弈为两人零和博弈。这是因为此种情况的盈利矩阵减去一个常数盈利矩阵（即矩阵所相应的博弈中所有局中人在所有结局中的盈利均为常数 C）后所得到的仍是两人零和博弈的盈利矩阵。而所谓的常数盈利矩阵表明局中人无论采取什么策略，他们的得失总是一样，因此在对应常数盈利矩阵的博弈中无所谓去预测局中人将采取何种策略。于是，每个结局里局中人盈利之和等于常数的两人博弈的预测结果与除去常数盈利矩阵之后的两人零和博弈的预测结果完全一样。

不管博弈的最终结局如何，两人零和博弈中的两个局中人总

有一方赢而另一方输(且输赢数量相等)。应当承认这是博弈论中一个比较极端的情况,但在社会经济活动中有时会出现类似的两人零和博弈。例如最简单的股票交易,甲抛出100手某上市公司股票,被乙买进,倘若不计股票交易中必须的但相对来说却是极少量的印花税与手续费等费用,那么此类交易可以近似地视作两人零和博弈。可是在社会经济领域中,绝大部分令人感兴趣的博弈呈现非零和现象。两个国家之间正常的贸易就是一个非零和两人博弈,一般地他们通过双方之间的贸易达到互惠互利,否则贸易不可能年复一年地继续下去,每一次贸易或每一次博弈中的每一个结局常常是"正数和"盈利。

尽管猜谜博弈由于它的零和盈利而极具特殊性,但是它的表示形式却在一类博弈的表示形式中具有代表性,因为这类表示形式的共同特点是给出了(1)博弈的局中人集合,(2)每个局中人的纯策略空间,以及(3)每一个局中人在每一个可能结局上的盈利。我们称此类表示型式为博弈的策略型表示,也称作博弈的正则型表示。为了体现普遍性,我们在以下正式定义中将两人博弈推广到n人博弈:

定义 2.1 n人博弈的正则型(或策略型)表示指定了n个局中人以及他们各自的纯策略空间$S_1,S_2,\cdots,S_n$和这些局中人各自得到的盈利函数$u_1,u_2,\cdots,u_n$。我们将该博弈表示为:$G=\{S_1,S_2,\cdots,S_n;u_1,u_2,\cdots,u_n\}$。

我们将在以下几节对猜谜这一策略型博弈进行进一步的讨论。

§2.2 混合策略

一个人倘若参加猜谜博弈,无论他处于甲的地位还是处于乙

的地位，只要他是理性的，他当然想赢，因此他很想知道他究竟是伸出一个指头较好还是伸出两个指头比较妥当，或者问，他应当采取哪一种策略呢?只要仔细研究一下猜谜博弈模型，简单的逻辑推理立即指出，不管局中人(甲或者乙)采取何种纯策略，他都有可能赢钱，也有可能输钱。除了存在作弊现象外，没有一个人会告诉局中人应该伸出几个手指。说得“学究气”一点，猜谜博弈在纯策略范围内不存在解。

当然，不存在纯策略解，不等于局中人不玩猜谜游戏。由于输赢的刺激，也许甲与乙会乐意猜谜多次。在每一次操作中，为了赢钱，任何一个局中人都不愿意把自己选择何种纯策略的意图暴露给对方，最好的办法就是从两个纯策略中随机地选择一个。也就是说，局中人1(甲)对于两个纯策略(伸一个指头，伸两个指头)各赋予一定的概率 p_1 与 p_2，显然 $p_1 \geqslant 0$，$p_2 \geqslant 0$，且 $p_1 + p_2 = 1$。向量 $p = (p_1, p_2)$ 实际上就是甲在猜谜博弈中采取的策略，它把两个纯策略用概率向量 p“混合”在一起考虑，因此称 p 为甲的混合策略。相应地乙也有他自己的混合策略 $q = (q_1, q_2)$。

混合策略的引进，使局中人“忽然”有了无穷多个策略。其实，原先的两个纯策略本身就是特殊的混合策略，只不过它们分别对应于退化的概率分布(1,0)与(0,1)。众多的混合策略有着各不相同的效果，以怎样的概率分布作为局中人选择合理的混合策略，相当于既然在猜谜博弈中没有纯策略解，那么不妨退一步看看混合策略有没有解。对于局中人1来说，如果混合策略 $(p, 1-p)$ 比较合适，那么从输赢的角度来看，无论局中人2采取什么策略，至少不能让他(局中人2)赢钱。但是，由于偶然性，在一次博弈中没有把握一定能办到这件事。设想如果博弈进行多次，局中人1取 $(p, 1-p)$，至少从平均的角度可以不让局中人2赢。或者说，不管局中人2出示几个手指，应使局中人2的平均(或期望)盈利不会大于零。用数学公式表示为：

局中人 2 若出示一个手指，期望盈利为

$$-p+(1-p)=1-2p\leqslant 0 \quad (2.1)$$

局中人 2 若出示两个手指，期望盈利为

$$p-(1-p)=2p-1\leqslant 0 \quad (2.2)$$

要使(2.1)式与(2.2)式同时成立，当且仅当 $p=1/2$，局中人 1(甲)的理想的混合策略应为(1/2,1/2)。同理，我们也可以求得局中人 2 的理想混合策略，也为(1/2,1/2)。

可能有人会争辩，博弈的最终结局一定表现为纯策略组合，绝对不会呈现概率向量的形式。甲、乙二人理想的混合策略组合((1/2,1/2),(1/2,1/2))蕴含着在每次猜谜之前，他们二人都要背过身去独立地各自抛一枚均匀的钱币，正面朝上则出一个指头，反面朝上则出两个指头。也许世界上猜谜的人还没有傻(或"机械")到这种地步！其实，倘若猜谜博弈一遍又一遍地重复许多次，两位局中人每次独立地从两个纯策略中等可能地抽取一个作为该次的行动，那么从平均的意义上来说，谁也不输谁也不赢，这不能不说是一个"皆大欢喜"的结果。我们称如此博弈的数值是零！

现在概括一下混合策略，先仅限于纯策略空间是有限维情况。

局中人 $i(i=1,2,\cdots,I)$ 的一个混合策略是该局中人的纯策略空间 $S_i=(S_{i1},S_{i2},\cdots,S_{ik_i})$($S_i$ 共有 k_i 个纯策略)上的概率分布，以数学符号 σ_i 来表示。所有 I 个局中人各自采取的混合策略 $\sigma_1,\sigma_2,\cdots,\sigma_I$ 是统计独立的。以这 I 个混合策略构成的"形式上的向量" $\sigma=(\sigma_1,\sigma_2,\cdots,\sigma_I)$ 实质上是博弈各方采取的策略"组合"，由于诸 σ_i 是概率分布，我们称这个策略"组合"为一个策略剖面(profile)。局中人 i 在该策略剖面上的盈利是该剖面上所有可能的纯策略组合盈利的期望值。下面我们进行具体的计算：

局中人 i 的混合策略全体可以表示为

$$\sum\nolimits_i=\{\sigma_i\} \quad (2.3)$$

记号 $\sigma_i(s_{ij})$表示概率分布 σ_i 赋予纯策略 s_{ij}的概率。如果读者稍有些数学常识,那么容易知道混合策略剖面的全体 $\sum=\{\sigma\}$实际上是 I 个 $\sum_i\ (i=1,2,\cdots,I)$的乘积空间:

$$\sum=\{\sigma\}=X\{\sum_i\}=\{(\sigma_1,\sigma_2,\cdots,\sigma_I)\} \tag{2.4}$$

假如我们知道局中人 i 取定纯策略 s_{ij}时的盈利为 $u_i(s_{ij})$(由于其他局中人采取的是混合策略,无疑 $u_i(s_{ij})$ 实质上也是期望盈利),那么局中人 i 在混合策略剖面 σ 上的期望盈利应当等于

$$u_i(\sigma)=\sum_{j=1}^{k_i}\sigma_i(s_{ij})u_i(s_{ij}) \tag{2.5}$$

关键在于 $u_i(s_{ij})$如何计算。我们知道,尽管 σ 是混合策略剖面,它的实现一定是 I 个局中人各从自己的纯策略空间中取某一个纯策略的组合,这样的纯策略组合的可能数目显然有 $k_1\cdot k_2\cdot\cdots\cdot k_I$ 个。倘若局中人 i 取定纯策略 s_{ij},与此匹配的纯策略组合可能个数为 $k_1\cdot\cdots\cdot k_{i-1}\cdot k_{i+1}\cdot\cdots\cdot k_I$。这些结局中的每一个的盈利不妨记作

$$\left.\begin{array}{l} U_i(S_{ij,t_1,\cdots,t_{i-1},t_{i+1},\cdots,t_I}) \\ t_1=1,2,\cdots,k_1;\cdots;t_{i-1}=1,2,\cdots,k_{i-1};t_{i+1} \\ \quad=1,2,\cdots,k_{i+1};\cdots;t_I=1,2,\cdots,k_I \end{array}\right\} \tag{2.6}$$

注意到 $\sigma_i(i=1,2,\cdots,I)$的统计独立假设,在给定局中人 i 取定 s_{ij} 的条件下每个结局 $S_{ij,t_1,\cdots,t_{i-1},t_{i+1},\cdots,t_I}$发生的条件概率为

$$\begin{aligned} &\sigma_1(S_{1t_1})\sigma_2(S_{2t_2})\cdots\sigma_{i-1}(S_{i-1,t_{i-1}})\sigma_{i+1}(S_{i+1,t_{i+1}})\cdots\sigma_I(S_{I,t_I}) \\ &\triangleq\sigma_1(t_1)\sigma_2(t_2)\cdots\sigma_{i-1}(t_{i-1})\sigma_{i+1}(t_{i+1})\cdots\sigma_I(t_I) \end{aligned} \tag{2.7}$$

(2.7)式第二行的表示是不规范的,但书写很方便。因此局中人 i 在混合策略剖面 σ 上的盈利函数等于

$$\begin{aligned} u_i(\sigma)=&\sum_{j=1}^{k_i}\sigma_i(s_{ij})\sum_{t_1=1}^{k_1}\cdots\sum_{t_{i-1}=1}^{k_{i-1}}\sum_{t_{i+1}=1}^{k_{i+1}}\cdots\sum_{t_I=1}^{k_I}\sigma_1(t_1)\cdots\sigma_{i-1}(t_{i-1})\sigma_{i+1}(t_{i+1})\cdots\sigma_I(t_I) \\ &\cdot u_i(s_{ij,t_1,\cdots,t_{i-1},t_{i+1},\cdots,t_I}) \end{aligned}$$

$$=\sum_{j=1}^{k_i}\sum_{t_1=1}^{k_1}\cdots\sum_{t_{i-1}=1}^{k_{i-1}}\sum_{t_{i+1}=1}^{k_{i+1}}\cdots\sum_{t_I=1}^{k_I}\sigma_i(s_{ij})\sigma_1(s_{1_{t_1}})\cdots\sigma_{i-1}(s_{i-1,t_{i-1}})\sigma_{i+1}(s_{i+1,t_{i+1}})$$
$$\cdots\sigma_I(s_{I,t_I})\cdot u_i(s_{1,t_1},\cdots,s_{i-1,t_{i-1}},s_{i+1,t_{i+1}},\cdots,s_{I,t_I}) \tag{2.8}$$

从理论上说，$u_i(\sigma)$的计算——即(2.8)式——是重要的，每一个局中人在作出其决策之前必须比较一下各种可能结局上盈利的多少。但是对于数学知识有些欠缺的人，(2.8)式实在太复杂了。因此一般博弈论教科书不给出(2.8)式而常代之以一些形似“粗糙”的数学公式。其实，即使对于那些数学较好的人，也未必需要生搬硬套公式(2.8)式去计算期望盈利。一般仅需从实际情况出发依局中人 i 取定每一 s_{ij}时求期望盈利而逐一完成计算。下面以一个简单的两人博弈为例：

例 2.1　二人博弈的盈利矩阵(如图 2.2 所示)。

局中人 2

		L	M	R
局中人 1	U	4,3	5,1	6,2
	M	2,1	8,4	3,6
	D	3,0	9,6	2,8

图 2.2

若局中人 1 的混合策略取为

$$(\sigma_1(U),\sigma_1(M),\sigma_1(D))=(\frac{1}{3},\frac{1}{3},\frac{1}{3})$$

而局中人 2 的混合策略取为

$$(\sigma_2(L),\sigma_2(M),\sigma_2(R))=(0,\frac{1}{2},\frac{1}{2})$$

那么在混合策略剖面(σ_1,σ_2)下局中人 1 的期望盈利为：

$$u_1(\sigma_1,\sigma_2)=\frac{1}{3}(0\times4+\frac{1}{2}\times5+\frac{1}{2}\times6)$$

(局中人 1 取 U 时的期望盈利$\times\sigma_1(U)$)

$$+\frac{1}{3}(0\times2+\frac{1}{2}\times8+\frac{1}{2}\times3)$$
(局中人 1 取 M 时的期望盈利$\times\sigma_1(M)$)
$$+\frac{1}{3}(0\times3+\frac{1}{2}\times9+\frac{1}{2}\times2)$$
(局中人 1 取 D 时的期望盈利$\times\sigma_1(D)$)
$$=\frac{11}{6}$$
类似地，我们可以求得该情况下局中人 2 的期望盈利为：
$$u_2(\sigma_1,\sigma_2)=0+\frac{1}{2}\left(\frac{1}{3}\times1+\frac{1}{3}\times4+\frac{1}{3}\times6\right)$$
$$+\frac{1}{2}\left(\frac{1}{3}\times2+\frac{1}{3}\times6+\frac{1}{3}\times8\right)$$
$$=\frac{27}{6}$$

§2.3 累次严优

在博弈中局中人在选择自己的决策时要比较各种策略的盈利大小，哪一个策略所获盈利大就选那个策略，可是就纯策略来说，即使是最简单的两人博弈，由于对手的纯策略空间通常不止一个元素，因此局中人在每一个纯策略的盈利一般不是一个数而是一个向量，以图 2.2 的博弈为例，局中人 1 取策略 U、M、D 时的盈利向量分别是(4,5,6)，(2,8,3)和(3,9,2)。显然，这三个向量无法比较大小，因此从盈利向量的角度出发，局中人 1 无法立即作出决策，他只有在预测对手的可能行动后才能下决心。譬如，假如他预测局中人 2 极可能采取策略 R，相应地局中人 1 取 L、M、D 的盈利分别为 6、3、2，于是局中人 1 会毫不犹豫地取策略 L。当然，预测是存在一定风险的。现在再从局中人 2 的角度出发考虑问题，不管局中人 1 如何行动，局中人 2 取策略 L、M、R 时的盈利向量分别

为(3,1,0),(1,4,6)和(2,6,8)。这三个向量中,第一个向量与后面两个向量无法比大小,但是第二个向量中的每个元素均小于第三个向量的相应位置元素,这表明,无论局中人1怎样决策,局中人2与其选取策略 M 还不如选取策略 R。对于局中人2来说,M 是个劣策略。所谓某局中人的劣策略,就是指在他的纯策略空间中至少存在一个其他的纯策略在盈利向量的每个元素优于该策略的盈利向量的相应元素(这种说法是对劣纯策略而言,对于混合策略的劣策略将在稍后正式提及),理性的局中人肯定不会采用劣策略。

如果博弈处于完全信息状态,理性的局中人1十分明白理性的局中人2不会采取劣策略 M,基于这样的认识,局中人1在分析与预测博弈的各种可能性时仅需从图2.3所示的盈利矩阵角度出发考虑自己的合适选择。

		局中人2	
		L	R
	U	4,3	6,2
局中人1	M	2,1	3,6
	D	3,0	2,8

图 2.3

在这个合理地被"缩小"了的博弈问题中,利用刚才引进的劣策略概念,对于局中人1而言,此时 M 与 D 成为劣策略(注意:在原来的博弈中,M、D 不是局中人1的劣策略),或者说,策略 U 优于 M 与 D。这里,M 与 D 是局中人1的"条件"劣策略,所谓条件就是局中人1预测局中人2必定舍弃自己的劣策略 M。此时的局中人1也必定舍弃条件劣策略 M、D 而选择 U。局中人2凭着自己的理性可以分析出局中人1会预测到自己将舍弃 M,同时他也必定可以预测到在缩小了的博弈中局中人1将舍弃 M 与 D。剩下两个可能结局:(U,L)——两人的盈利为(4,3);(U,R)——两人

的盈利是(6,2)。无可非议,局中人2只有选取策略L才有可能提高自己的盈利。于是,我们可以预测博弈的合理结局可能是(U,L)——局中人1取策略U,与此同时局中人2采用策略L。

上述讨论中,每一个局中人舍弃自己的劣策略或条件劣策略的做法从逻辑上说是令人信服的。从某一个局中人的角度出发排除该局中人的劣策略,然后在"缩小了的""条件"博弈中,从另一个局中人角度出发,剔除此人的劣策略,这样一步一步地进行下去,除非到某一步不存在所谓的劣策略,否则可以一直剔除下去,最后幸存下来的结局合乎逻辑地成为博弈的预测结局。我们称这个过程为累次取优(iterated dominance),更精确地,应称为累次严优(iterated strict dominance)。之所以要使用"严"字,是因为我们所剔除的策略必须是严劣的,就是说,至少存在一个纯策略其盈利向量的每一个元素严格地大于劣策略的盈利向量中的相应元素。

逻辑上的严密性使得人人都愿接受累次严优为求博弈问题解的好办法,但是它也留下了若干需要思考的问题:

(1)在图2.2所示的博弈中,已经发现局中人1不存在劣策略,我们是从局中人2的角度出发才开始累次严优过程的。倘若此时局中人2也不存在劣策略(这是可能的,例如在图2.2中,我们将(M,M)的相应盈利向量改为(8,7),此时无论L、M或者R都不是局中人2的劣策略。)显然在这种场合无法使用累次严优法,然而仍需进行预测。可见,累次严优法对于预测博弈的合理结局是有局限性的。

(2)如果在图2.2所示博弈中,我们将(M,M)的盈利改为(3,4)。此时局中人1在纯策略M的盈利向量为(2,3,3),它的每个元素均小于局中人1在纯策略U的盈利向量(4,5,6)的相应部分。就是说,此时M是局中人1的严劣策略,而局中人2的严劣策略仍为M。令人忧虑的问题是累次剔除劣策略过程先从局中人1开始与先从局中人2的角度出发其最后结果是否一样?倘若结果

不一样，那么累次严先似乎成为不可信任的方法。

(3)累次严优过程是在比较了纯策略的盈利向量之后再剔除严劣纯策略的。我们已经知道纯策略空间有必要推广到混合策略空间，因此一个纯策略如果是严劣的话，照理应当与混合策略空间中的任意策略作比较。我们面临的问题是在纯策略之间无优劣之分时，是否可以考虑混合策略的优劣，或者问混合策略是否有严劣策略之说，更要问，是否可能一个纯策略可以不劣于任何其他纯策略，但是它却劣于某些混合策略从而列入被剔除名单，等等。

我们将逐一对上述问题作出较通俗的回答，顺便对劣策略用数学语言给予严格的描述。

第一个问题已经暴露了累次严优解的局限性。由于累次剔除严劣策略过程中逻辑上的严密性，必然大大缩小预测合理结局的范围，因此有必要将解的范围进行拓广，这个问题在稍后介绍 Nash 均衡概念时将得到解决。

累次严优解的范围窄小常常使得它成为博弈问题的“理想”预测，但是必须指出的是，有时候理想的事情未必在实际生活中行得通，尤其是在盈利函数取极端值时会出现“反常”现象。比如下面一个简单的两人博弈(见图 2.4)：

局中人 2

		L	R
局中人 1	U	7,9	−1 000,8.5
	D	6,5	5,4.5

图 2.4

从局中人 2 的观点，显然 L 优于 R，即 R 是局中人 2 的严劣策略，在完全信息情况，局中入 1 预测局中人 2 会剔除 R，在(U,L)与(D,L)两个可能结局中，他自然选取策略 U，于是累次严优过程产生了唯一解(U,L)。然而，预测并不等于必然发生，局中人

1 预测局中人 2 会舍弃策略 R，并非表示局中人 2 毫无问题地选取策略 L。从图 2.4 不难看出，不管局中人 1 如何行动，局中人 2 取 R 仅比取 L 在盈利上差 0.5——一个微不足道的量。因而很有可能局中人 2 对在 L 与 R 之间如何选择表现出无所谓态度。倘若这样，局中人 1 可能应当意识到此时取策略 D 也许比取 U 更稳健一些。一旦局中人 2 采取了策略 R(纵然是千分之一的可能性)，局中人 1 不至于蒙受必须付出1 000的巨大损失。而即使局中人 2 的确舍弃了 R，局中人 1 取 D 也比 U 只少了盈利 1。我们称这样分析预测博弈的思路具有稳健性，它类似于统计学中 Robustness 思想。稳健地预测博弈达到合理可能结局在社会经济领域中具有重大的意义，它可以使局中人避免过大的风险。因此，在现代博弈论研究中，"稳健性"成为"精炼"均衡的要求之一。稳健性考虑显然取决于盈利函数，假如把盈利矩阵中的－1 000改成 6 或 5 之类，显然累次严优立刻显示出它的优越性与合理性。

至于第二个问题，其提出本属杞人忧天。从不同局中人观点出发的累次严优过程，只要有可能实施，其最后幸存者必定是同一个结局。该结论主要得益于严劣策略的定义，假如某局中人有一个严劣策略，那么这个策略的盈利向量的任何子向量(即去除某些分量后剩余的分量所组成的向量)都将劣于那个严优于它的纯策略的盈利向量的相应子向量。因此即使先从其他局中人角度出发开始累次严优过程，在经过剔除后剩余下来的"缩小了的"、"有条件的"博弈中局中人将毫不犹豫地舍弃这个原先从他那儿开始本应剔除的严劣策略。用一句话来概括，只要是严优策略，无论先从哪个局中人的角度出发，在累次严优过程中它逃脱不了被剔除的命运。

关于第三个问题的讨论，有必要将劣策略概念从纯策略范围解脱出来，我们用数学语言在较严格的意义下刻画策略的优劣。

博弈论常关心当对手确定策略时局中人 i 的策略变化。设有 I 个局中人，若局中人 i 的策略如前简记为 $s_i \in S_i$(S_i 为局中人 i 的纯

策略空间)，除 s_i 之外由其他 $(I-1)$ 个局中人采取的策略向量简记为 s_{-i}，它所来自的乘积空间简记为 S_{-i}，$s_{-i}\in S_{-i}$。于是 I 个局中人的纯策略组合向量 $(s_1,\cdots,s_{i-1},s_i,s_{i+1},\cdots,s_I)$ 可以简记为 (s_i,s_{-i})，将这些方便的记号拓广到混合策略，从而混合策略剖面 $(\sigma_1,\cdots,\sigma_{i-1},\sigma_i,\sigma_{i+1},\cdots,\sigma_I)$ 简记作 (σ_i,σ_{-i})。相应于 (s_i,s_{-i}) 与 (σ_i,σ_{-i}) 的局中人 i 的盈利或期望盈利函数可以记为 $U_i(s_i,s_{-i})$ 与 $U_i(\sigma_i,\sigma_{-i})$。

定义 2.2 对于局中人 i 的(混合)策略空间 $\sum_i$ 中的某个纯策略 s_i，如果存在混合策略 $\sigma_i^*\in\sum_i$ 使得

$$U_i(\sigma_i^*,s_{-i})\geqslant U_i(s_i,s_{-i}) \tag{2.9}$$

对任意 $s_{-i}\in S_{-i}$ 成立，且在 S_{-i} 中至少存在一个纯策略组合 $s_{-i}^*\in S_{-i}$，使(2.9)式中的不等号严格成立

$$U_i(\sigma_i^*,s_{-i}^*)>U_i(s_i,s_{-i}^*) \tag{2.10}$$

则称纯策略 s_i 为局中人 i 的弱劣纯策略。倘若对一切 $s_{-i}\in S_{-i}$，(2.9)中的不等式都严格地成立

$$U_i(\sigma_i^*,s_{-i})>U_i(s_i,s_{-i}) \qquad \forall s_{-i}\in s_{-I} \tag{2.11}$$

则称 s_i 为局中人 i 的严劣纯策略。

定义 2.2 已经指出了局中人 i 的劣(或严劣)纯策略是通过与混合策略空间 $\sum_i$ 中的一切其他元素进行盈利大小的比较而定义的，因此纯策略与混合策略的优劣比较自然不成任何问题。但是定义 2.2 也给人带来悬念，因为在定义 2.2 的不等式中始终将对手的策略剖面限制在纯策略组合向量空间 S_i 中，依常理，策略的所谓优与劣，其不等式(2.9)式或(2.11)式应当对其他局中人的混合策略空间 $\sum_{-i}$ 中的任何元素 σ_{-i} 成立。那么定义 2.2 是否有悖于常理呢？为了弄清楚这个问题，不妨考虑最简单的两人博弈，即 $I=2$，并为方便起见，仅假设局中人 2 的纯策略空间中仅含 s_{21} 与 s_{22} 两个元素，局中人 2 的任意混合策略可记作 $\sigma_2=(p,1-p)$，其中 $0<p<1$。利用计算盈利函数的基本技巧，易得

$$u_1(\sigma_1,\sigma_2) = pu_1(\sigma_1,s_{21}) + (1-p)u_1(\sigma_1,s_{22}) \quad (2.12)$$

(2.12)式表明局中人 1 在混合策略剖面 (σ_1,σ_2)的盈利函数实际上是其对手取纯策略时盈利函数的凸组合，以严劣策略为例，倘若 s_{11}是局中人 1 的严劣策略，由定义 2.2 应有

$$u_1(\sigma_1,s_{21}) > u_1(s_{11},s_{21}),\ u_1(\sigma_1,s_{22}) > u_1(s_{11},s_{22})$$

(2.12)式蕴含了

$$\begin{aligned} u_1(\sigma_1,\sigma_2) &= pu_1(\sigma_1,s_{21}) + (1-p)u_1(\sigma_1,s_{22}) \\ &> pu_1(s_{11},s_{21}) + (1-p)u_1(s_{11},s_{22}) = u_1(s_{11},\sigma_2) \end{aligned} \quad (2.13)$$

这说明既然(2.11)式对所有纯策略 s_{21}与 s_{22}成立，那么局中人 1 的盈利函数作为剖面中局中人 2 分别取 s_{21}与 s_{22}时局中人 1 相应盈利函数的凸组合也满足(2.11)式，于是(2.11)式对一切混合策略 $\sigma_2\in\sum_2$（即 $\sigma_{-1}\in\sum_{-1}$）一定成立。正因为如此，我们在定义 2.2 中只要考虑 S_{-i}就足够了。

定义 2.2 表明可以有纯策略劣于混合策略的现象，具体地我们仅用一个小例子就可以看到（见图 2.5）。

局中人 2

局中人 1		L	R
	U	2,0	−1,0
	M	0,0	0,0
	D	−1,0	2,0

图 2.5

图 2.5 所表示的博弈中，从(2，−1)、(0,0)及(−1,2)三个盈利向量可知局中人 1 的纯策略 U、M、D 三者之间无法比较优劣，似乎不存在严劣策略，假如局中人 1 取混合策略(1/2,0,1/2)——即以 1/2 概率取 U,0 概率取 M 及以 1/2 概率取 D，此时不管局中人 2 采取何种行动，不妨设他取混合策略$(p,1-p)$，局中人 1

的期望盈利为：

$$[2p-(1-p)]1/2+0+1/2[-p+2(1-p)]=1/2$$

它大于局中人 1 取纯策略 M 时的盈利 0（无论局中人 2 取 L 还是 R）。因此纯策略 M 是局中人 1 的严劣策略。对于局中人 1 来说，宁取混合策略（1/2，0，1/2）而不愿意取 M——这倒是符合一般赌徒的心理，与其坐在家里不输也不赢，不如冒些风险去赌一把。这个例子只是进一步阐明定义 2.2 中所指出的局中人的严劣策略是在该局中人混合策略范围内比较盈利多少而得出的。读者千万不要误以为混合策略一定不会差。其实只要存在严劣纯策略，那么任何赋予该严劣纯策略以正概率的混合策略一定也是个劣策略，因为将这点上的正概率添加到优于严劣纯策略的另一个纯策略去，从而得到的新的混合策略，其期望盈利一定会得到增加。也许又会有人提出问题，假如一个混合策略仅对那些非劣纯策略赋予正概率，那么它有否可能是个劣策略？答案是明白无误的，一个混合策略是一个劣策略，责任不能全怪罪于严劣纯策略的存在与作用。试看下述例子（见图 2.6）：

局中人 2

局中人 1		L	R
	U	1，3	−2，0
	M	−2，0	1，3
	D	0，1	0，1

图 2.6

观察局中人 1 的混合策略（1/2，1/2，0），若局中人 2 取策略 L，局中人 1 的期望盈利为

$$1/2\times1-2\times1/2=1/2-1=-1/2$$

如果局中人 2 取 R，此时局中人 1 的期望盈利是

$$-2\times1/2+1\times1/2=-1+1/2=-1/2$$

这两种情况的平均盈利均小于策略 D 时的盈利(0,0),因此混合策略(1/2,1/2,0)是个劣策略。局中人 1 不会采纳这个混合策略,但它并非是对严劣纯策略赋予正概率而得,因为这里 U 与 M 不是严劣策略。

§2.4 累次严优的应用实例

本节我们讨论若干经典的博弈问题,它们利用累次严优而得到合理预测,绝大多数教科书中可以见到它们或者类似的例子。在一定程度上它们展现了博弈论的魅力所在。

例 2.2 囚徒窘境(Prisoner's Dilemma)。

甲、乙两人共同作案而被警方抓获,他们分别被关在不同的屋子里接受警方的盘问,并获知警方向他们公开的有关政策(两个局中人的共同知识)。在互不知晓同伴怎样做的情况下,甲乙两人都面临两种选择——坦白与抗拒。如果甲乙双方互相合作而拒不坦白,那么由于证据不足从而在各关一个月后获得释放;如果他们都采取背叛对方的策略而坦白全部作案事实,那么根据案情将被囚禁 8 个月;如果两人中有一人拒不坦白,而另一个人却坦白罪行,那么坦白者因立功而被释放,抗拒者则受到严惩被关 15 个月。因此这构成一个完全信息的静态博弈。两个局中人的切身利益是关心囚禁多少时候而不是盈利多少,因而我们称图 2.7 所示的矩阵为效用矩阵(payoff matrices)。

		乙 坦白	乙 抗拒
甲	坦白	−8,−8	0,−15
甲	抗拒	−15,0	−1,−1

图 2.7

图 2.7 中效用矩阵对于甲、乙两人是对称的,因此我们仅考虑

从甲的角度出发，甲坦白的效用向量(—8，0)显然大于甲抗拒的效用(—15，—1)(因为—8>—15，0>—1)，因此抗拒是甲的严劣策略，由对称性，抗拒也是乙的严劣策略。理性的甲与乙都会舍弃劣策略，于是(坦白，坦白)成为博弈的累次严优解。博弈双方为了自身利益导致了(—8，—8)效用的结局，“可惜”这是一个“无效”的结局。

这里我们引用了经济学中“有效”的术语。如果不存在其他的结局，使得某些局中人的效用(或盈利)比在这个结局的效用好得多，同时又不会使其他局中人的效用(或盈利)变得更差，则称博弈的这个结局是有效的(efficient)。反过来，如果一个结局不是有效的，则一定(至少)存在另外一个受到局中人一致欢迎的结局。

显然，在囚徒窘境模型中，(坦白，坦白)不是有效的，因为从效用角度出发，(抗拒，抗拒)对两个囚徒更具有吸引力。(抗拒，抗拒)是有效的结局，但是它不是博弈的解。这正是博弈论引起人们兴趣的地方。它揭示了局中人理性行为的结局可以不是有效的。

囚徒窘境博弈模型告诉我们，在完全信息且静态的情况下，为了自己切身利益，两个囚徒都愿意走“坦白从宽”道路。然而，无论是中国的公安机关，还是美国的警察局，警方也许会告诉你，在两人共同作案的案件中，互相背叛而坦述案情所占的比例并不令人满意。这是因为对于两个囚徒来说，这样的博弈不是一次性的，常常是重复多次，存在着动态博弈，这中间存在“威胁”与“承诺”等因素，使得效用发生根本变化。关于动态博弈、重复博弈的研究将在本书以后几章中逐步展开。

与囚徒窘境类似的博弈问题在社会经济领域内有着许许多多的版本。例如，两个公司竞争出售同一产品，可以有高与低两种价格，双方合作以高价格垄断市场，可以使大家获得满意的利润，至少远远好于双方都以低价格出售产品的情况。但是如果一方坚持高价而另一方为了独占市场将产品设置至低价格的话，那么后者

将获最高盈利而前者则损失惨重，显然这是一个类似囚徒窘境的博弈，其盈利矩阵几乎与囚徒窘境中的效用矩阵一样，除了具体数值有所变动。不言而喻，双方都以低价格出售商品是博弈的合理预测。市场上的价格大战颇能反映这个问题。

两个互通贸易的国家理性地做生意，从而互惠互利。然而各自的国家利益会驱使它们互设障碍，诸如提高关税、在议会通过反倾销法案等等。其实，如果它们愿意去除这些障碍的话，也许会给双方带来更大的利益。

买卖双方在市场上拼命地讨价还价，以至于无法达成一个双方都能接受的方案。如果双方真心诚意地合作，常常存在一个价格，使得卖方获得较为满意的利润而买方则以适当的价格得到自己需要的东西。

下面以一个典型的例子结束“囚徒窘境”的讨论，那就是苏联与美国之间的扩军与裁军问题。

大多数善良的人会认为，如果世界上每个国家都将钱用于和平建设而不搞军备竞赛，无疑对人类是个福音，无论是穷人还是富人、穷国还是富国，都必将处境更好。但是现实告诉人们，适当的军备可能是必要的。让我们用一个博弈模型解释这个矛盾。

假定世界上仅有苏联与美国（这个假设似乎离现实太远，但在当时，这两个所谓超级大国的确并没有将其他国家放在眼里），它们都面临着在两个策略之间进行选择：扩充军备（即武装）或者裁军。如果双方都热衷于军备竞赛，必将为此付出3 000亿美元（数字是虚拟的）代价；如若双方裁军，则可省下这笔钱，即支付下降到零。但是倘若有一方裁军而另一方进行武装的话，武装的一方发动侵略占领对方国土从而获益 1 万亿美元，裁军的一方由于军事失败而丧失国土从而可认为损失是无限的。图 2.8 列出的获益矩阵除了具体数字外，基本上等同于囚徒窘境的效用矩阵。累次剔除严劣策略过程告诉我们（武装，武装）是博弈的解，尽管（裁军，裁军）

是有效的结局。

		苏联	
		武装	裁军
美国	武装	−3 000,−3 000	10 000,−∞
	裁军	−∞,10 000	0,0

图 2.8

如果双方在战争与和平这个问题上都能理智一些的话(注意,这里的理智完全不同于博弈论中局中人的理性),其实大家的日子都会过得好一些。然而,博弈论却预测双方武装可以达到稳定的平衡。一个真正荒谬但却符合现实的合理结局,使人们对博弈论产生浓厚兴趣。

例 2.3 智猪博弈。

大猪与小猪喂养在同一个猪圈中,猪圈的一头安装有一杠杆,只要一踩杠杆,猪圈的另一头固有的食物槽里将会流出饲料。踩杠杆需要花费能量,相当于消耗半份饲料,大小猪都不踩的话,它们虽然不耗费热量但吃不到任何东西。设食物槽内一次流出的饲料共有 6 份,如果小猪踩杠杆,等它跑到食物槽跟前时,将发现不劳而获的大猪已经吃光了全部 6 份饲料;而若大猪踩杠杆后再跑到食物槽跟前时,它将发现自己只能分享到 1 份饲料,其余的全被等候在那儿的小猪享用完了;倘若它们共同踩杠杆再一起跑向食物槽,则大猪能争得 4 份饲料而小猪只能争得 2 份饲料。将这些信息(这是大猪小猪的共同知识)归结为如图 2.9 所示的智猪博弈的盈利矩阵。

		大猪	
		踩	不踩
小猪	踩	1.5,3.5	−0.5,6
	不踩	5,0.5	0,0

图 2.9

小猪具有严优策略——不踩杠杆而坐享其成。在这种情况下,

大猪倘若不踩，那么它们都将饿肚子。如果大猪去踩杠杆，虽然让小猪占到了便宜，但至少自己还能尝到一份，比起踩杠杆所消耗的半份饲料的热量，毕竟有所得。因此智猪争食博弈模型的累次严优解是(大猪踩，小猪不踩)。

"囚徒窘境"与"智猪争食"虽然都可以通过累次严优过程获得博弈解，但它们之间存在着内涵的不同。囚徒窘境的一个鲜明特点是在"个别理性"与"共同理性"之间存在着矛盾。如果博弈双方能成功地合作，则双方的处境会好一些，可是在博弈中个别理性占了上风从而他们不愿合作，囚徒窘境的社会经济"版本"充分地体现了这个矛盾。智猪博弈不产生个别理性与共同理性之间的冲突。博弈的解(大猪踩，小猪不踩)是个有效结局，因为任何其他结局都将使小猪的收益减少。

用智猪博弈模型解释经济问题，最成功的例子之一是 OPEC (石油输出国组织)的分配方案。OPEC 的成功之处可能相当程度上归属于它的最大成员国——沙特阿拉伯的愿望。沙特希望所有的成员国都能节制石油产量以使油价保持在较高水平之上。当某些小石油输出国"偷偷"增加自己的产量时，沙特阿拉伯"大度地"削减自己的产量以保持总产量的稳定。在这里沙特阿拉伯扮演了"大猪"的角色。因为沙特阿拉伯与那些小石油输出国都明白此时除非沙特限制自己的产量，否则 OPEC 可能面临崩溃；小成员国依赖于沙特阿拉伯对 OPEC 的努力而从中渔利。事实上，沙特阿拉伯为了自己赢得高价利润的足够享受，理性地愿意忍受维持 OPEC 的不匀称摊派。

§2.5 Nash 均衡

如前所述，累次严优不失为预测博弈的好办法，不幸的是，经济博弈中不少问题未必可以通过累次严优求得解，尤其是在博弈

各方的纯策略空间中并不存在严劣纯策略的时候。本节的目的在于拓广累次严优解的范围，以使更广泛一类的博弈问题存在合理解。为此，我们分析累次严优过程之所以适用范围窄小的原因，是因为对局中人的严劣纯策略的定义要求过严，它需要在该局中人的策略空间中至少存在一个策略(可能是混合策略)，在给定其他局中人的每一可能策略(剖面)时均在盈利或效用上优于它。尤其在实施累次剔除严劣策略时，我们总是先比较两个纯策略的盈利(向量)大小。事实上，在一般的经济博弈中，局中人的任意两个纯策略，比如记作 A 与 B，常常在对手采取某种策略时，A 优于 B，而在对手取另一策略时，A 又劣于 B，不一定出现不管对手如何行动，A 总是优于 B 或者 B 总是优于 A 的现象。这个事实增加了博弈与预测博弈的复杂性，同时限制了累次严优的可行性。对于博弈中的每一个局中人，真正成功的措施应该是针对其他局中人所采取的每次行动，相应地采取最有利于自己的策略，或者说，对于对手的每一行为作出最有利于自己的反应。于是，每一个局中人应采取的策略必定是他对于其他局中人策略的预测的最佳反应。Nash 均衡正是体现出博弈的这一基本原则。

Nash 均衡因数学家 Nash 而命名，Nash 均衡策略是指这样的策略组合(或剖面)，为了极大化自己的盈利(或效用)，每一个局中人所采取的策略一定应该是关于其他局中人所取策略的最佳反应。因此没有一个局中人会轻率地偏离这个策略组合(或剖面)而使自己蒙受损失。用数学语言对上述想法概而括之，我们有如下定义：

定义 2.3 完全信息静态博弈问题中的混合策略剖面 σ^*，如果对所有局中人 $i(i=1,2,\cdots,I)$ 均成立

$$u_i(\sigma_i^*,\sigma_{-i}^*)\geqslant u_i(s_{ij},\sigma_{-i}^*)\quad \forall s_{ij}\in S_i \tag{2.14}$$

那么，σ_i^* 被称作该博弈的 Nash 均衡。

如果 σ_i^* 是退化的混合策略(也就是说，在纯策略空间上的概率分布是退化分布)，那么我们得到纯策略 Nash 均衡。细心的读者也许会指出，定义 2.3 定义了混合策略 Nash 均衡，照理(2.14)式应当对局中人 i 的混合策略空间 $\sum_i$ 中的一切 σ_i(可以包括 σ_i^*)成立，但定义中只要求局中人 i 的纯策略空间中的每一个 s_{ij} 成立(2.14)式即可。这种做法是否有局限性？回答是否定的，我们只要注意到局中人的期望盈利函数是关于他取纯策略时盈利函数的凸组合这一事实，就明白定义 2.3 的合理性。

Nash 均衡的引入与定义需要进一步说明一些问题：

(1)定义 2.3 所确定的 Nash 均衡是否如我们原先设想的那样是从累次严优解延拓而得的？

为回答这个问题，仅需说明，如果存在累次严优解的话，它必然也是 Nash 均衡。或者说，Nash 均衡包含了博弈的累次严优解。其实这是毫无疑问的，设 $s^*=(s_1^*,\cdots,s_I^*)$ 是剔除严劣策略过程后幸存的唯一策略组合，根据剔除严劣策略的法则，对于任意 $s_i(\neq s_i^*)$，在给定 s_{-i}^* 的条件下，它必定(至少条件地)严劣于 s_i^*，特别地

$$u_i(s_i^*,s_{-i}^*)\geqslant u_i(s_i,s_{-i}^*) \qquad \forall s_i\in S_i \quad s_i\neq s_i^*$$

即 s^* 满足(2.14)式，故它是 Nash 均衡，其实很容易明白它是真正的唯一 Nash 均衡。囚徒窘境博弈问题中，甲与乙两人均取不合作态度，即取(坦白，坦白)。它是累次严优解，因此是唯一的(纯策略的)Nash 均衡。

(2)Nash 均衡使得每一个局中人在给定对手策略时作出最佳反应，那么这个反应策略决不会是严劣纯策略(假如该局中人的纯策略空间中存在严劣纯策略的话)，而且任意混合策略 Nash 均衡必定仅在非严劣策略置于正概率。但是对于弱劣策略未必如此，其理由是因为定义 2.3 中的不等式不是严格的，它允许等号成立的可能。下面我们就用一个简单的例子来说明这一问题(见图 2.10)。

局中人 2

		L	R
局中人 1	U	3,0	2,2
	D	3,2	0,1

图 2.10

由 Nash 均衡定义容易验证(D,L)是 Nash 均衡解，但是不难看出，对于局中人 1 来说，策略 D 是弱劣策略，这表明 Nash 均衡可以发生在某局中人取弱劣策略的结局。我们清楚地看到由于 D 是弱劣策略，尽管当局中人 2 取定 R 时，局中人 1 取 D 的盈利小于他取 U 的盈利，但当局中人 2 取策略 L 时，局中人 1 取 U 与 D 可获同样大小盈利。也就是讲，当局中人 2 取策略 L 时，局中人 1 的最佳反应可以不只是一个。因此他在选取 U 还是 D 方面会表示出无所谓态度，这给预测的稳健性带来了挑战。因为若以 Nash 均衡(D,L)作为预测的话，人们不禁会惊异甚至产生怀疑，在局中人 2 取 L 时，凭什么局中人 1 非要去取 D 呢？

1973 年由 Harsanyi 引进的严格均衡概念就是针对上述情况的，它在某种场合似乎更激起人们的兴趣。

定义 2.4 在策略型博弈 $G(S_1,S_2,\cdots,S_n,u_1,u_2,\cdots,u_n)$中，如果每一个局中人关于其他局中人的策略具有唯一的最佳反应，这样的 Nash 均衡称作严格的(strict)。就是说，假如 $s^*=(s_i^*,s_{-i}^*)$是严格均衡，当且仅当 s^* 是 Nash 均衡。且对所有的纯策略 $s_i\neq s_i^*$，成立

$$u_i(s_i^*,s_{-i}^*)>u_i(s_i,s_{-i}^*) \tag{2.15}$$

由定义 2.4 可知，严格均衡必定是纯策略 Nash 均衡。但反过来，纯策略均衡未必是严格的，图 2.10 中(D,L)就是一个例子。严格均衡定义中的“唯一最佳反应”理所当然地受到从事应用的人们

的普遍欢迎，因为“唯一”二字使人们无须作其他考虑，为寻求博弈解提供极大方便，也增强了人们的“放心”程度。由于(2.15)式是严格不等式，倘若盈利函数稍微受到一点干扰，也不致影响(2.15)式的成立，原来的严格均衡在“新”的盈利函数下仍然是严格均衡。我们称严格均衡对于博弈的性能状态发生各种各样微小变化时具有稳健性。严格均衡的这些优点使得它更令人信服，更激起人们的兴趣，也使人们更乐于用严格均衡来预测博弈。可惜，严格均衡并非对所有的博弈问题都存在。最简单的例子莫过于猜谜博弈，在那里唯一的均衡是混合策略((1/2,1/2),(1/2,1/2))。

累次严优、严格均衡都比较理想，然而有时“可望而不可及”。其实在通常情况下，Nash 均衡还是比较能令人满意的，它在如下的意义下是如何去进行博弈的一个“相容”(consistent)预测：如果所有的局中人都预测某个 Nash 均衡将会发生，那么没有一个局中人会有兴趣或者会有激情去故意违背。Nash 均衡，也只有 Nash 均衡，才具有这样的良好性质：局中人可以预测它，也可以预测到他的对手会预测它的发生。假如发生了预测一个非 Nash 均衡的结局，那就意味着其中至少有一个局中人出了差错，要么在预测对手的行动时出错，要么在预测自身的最优盈利方面出错。当然谁也不能担保不会发生这类错误，譬如有时候“当局者迷，旁观者清”，局外人对博弈可能结局的情况了解多于局中人，这种场合局中人很有可能会发生预测差错。关于这种说法有如下两点补充：(1)不能因为无人能保证不出差错而舍弃 Nash 均衡，在大多数经济博弈应用中，人们还是常把注意力集中于 Nash 均衡；(2)Nash 均衡尽管是如何进行博弈的相容预测，并不等于它一定是个好的预测结局。因为事实上有时候博弈的可能结局依赖于更多的信息，它们比起单单由策略形式所提供的信息要多得多。例如，人们希望知道博弈局中人拥有多少经验，各自的文化背景如何，各人的习惯如何，等等。这些信息对于诸如外贸的讨价还价博弈是极为重要的，

如果人们只是学究式地考虑任何预测博弈，也许是草率的。

不管怎样，在社会经济领域的应用中，对于局中人，至少从理论上来说，只要 Nash 均衡存在，大多数场合下它是博弈如何进行的合理预测之一。接下来，我们探讨一下如何寻求 Nash 均衡解。通常总是先从定义出发。

不妨从简单的 2×2 博弈模型着手。设博弈有甲乙两个局中人，甲有两个纯策略——U 与 D，乙的纯策略空间也有两种选择——L 与 R。其盈利矩阵如图 2.11 所示。

		乙	
		L	R
甲	U	$\underline{a},\underline{e}$	$\underline{b},f$
	D	c,g	$d,\underline{h}$

图 2.11

设乙取定 L 时，甲取 U 时盈利为 a，取 D 时盈利为 c，此时甲的最佳反应仅需比较 a 与 c 的大小，不妨假设 $a>c$，即甲在乙取策略 L 时的最佳反应是 U，我们在 a 的下面划一短线作为记号。同样在固定乙取 R 时，比较 b 与 d 的大小，倘若 $b>d$，也在 b 的下面划一短线。调换一个角度，若甲取定 U，乙取何种策略为最佳反应呢？只需比较 e 与 f 的大小，假设 $e>f$，那么我们在 e 的下方划一短线，在甲取定 D 时，在 g 与 h 的较大者 h(假设)下面划一短线。图 2.11 中已经按假设划上了所有该划的短线，这说明相对于 L，甲取 U 为最佳反应，相对于 U，乙取 L 为最佳反应，依定义，(U,L) 为 Nash 均衡。事实上，这种划短线的方法告诉我们，(U,L) 为 Nash 均衡，当且仅当 a 与 e 下面同时划有短线，即当且仅当 $a\geq c$，$e\geq f$(在等号成立时两个盈利下均划短线)同时成立。注意到一个有趣的事实，a 不必与 d 作比较，e 也不必与 h 作比较，简言之，(U,L) 是否 Nash 均衡与结局 (D,R) 的盈利大小毫无关系。图

2.11 中的短线是由于我们对盈利数的若干假定之后才形成的，假如我们适当地调整关于各盈利数的假设，盈利矩阵可能成为如图 2.12、图 2.13 的形式或类似图 2.12、图 2.13 的形式。它们的共同特点是四个可能结局中没有一个出现双划线现象，依定义，该博弈没有纯策略 Nash 均衡解(当然并不排斥存在混合策略 Nash 均衡的可能)。不难发现，类似图 2.12 形式的盈利矩阵所表示的博弈问题中无论甲还是乙都没有劣纯策略。容易验证，在 2×2 博弈中，如果不存在纯策略 Nash 均衡，每个局中人的纯策略空间必定不存在劣纯策略。但反之不然，如果博弈存在 Nash 均衡，局中人未必有劣纯策略。图 2.14 所示的盈利矩阵就能说明这个问题。

$\underline{a},e$	$b,\underline{f}$
$c,\underline{g}$	$\underline{d},h$

图 2.12

或

$a,\underline{e}$	$\underline{b},f$
$\underline{c},g$	$d,\underline{h}$

图 2.13

$\underline{a},\underline{e}$	b,f
c,g	$\underline{d},\underline{h}$

图 2.14

易见，(U,L)与(D,R)均为 Nash 均衡，但两个局中人都没有严劣纯策略。

我们已经例举了 2×2 博弈问题可能不存在纯策略 Nash 均衡。在第三章中我们将证明博弈论中一个极其重要的结论——有限博弈必存在 Nash 均衡。这等于说，在类似图 2.12 那样的博弈问题中，还需要关心混合策略 Nash 均衡的寻求。在我们熟悉的博弈问题中，猜谜游戏也是一个不存在纯策略 Nash 均衡的重要例子，只要对它相应的盈利矩阵使用划线法就可以验证这一点。我们在讨论猜谜博弈时所求得混合策略解是不是 Nash 均衡呢？回答是肯定的。为说明这一点不妨研究猜谜博弈的“普及版本”——监

察博弈。这是一个广泛用于军备控制、纳税审查、制止犯罪等各个方面的博弈模型。

例 2.4 监察博弈。

代理商为委托人干活，有两个策略可供选择：工作(W)与偷懒(S)。倘若工作将使代理商花费 g，并理所当然地获得委托人付给他的工资 w ($w>g$ 是一个合理的假设，否则代理商没有任何激情去工作)。委托人在监督方面也有两个可供选择的纯策略：检查(I)与不检查(N)。如果委托人检查需要费用 h，以这点代价换得代理商是否在偷懒的信息。一旦发现代理商偷懒，则扣除工资作为惩罚，若代理商工作而不偷懒，则将为委托人增加价值 v 的财产(显然 $v>w$)。如果这些信息为共同知识，两个局中人进行完全信息静态博弈。为限制过多的各种应当考虑到的情况发生，不妨假定 $g>h>0$，其盈利矩阵如图 2.15 所示。

		委托人 I	委托人 N
代理商	s	$0, \underline{-h}$	$\underline{w}, -w$
	w	$\underline{w-g}, v-w-h$	$w-g, \underline{v-w}$

图 2.15

根据模型的假设，给盈利矩阵划上短线，如图 2.15 所示。现从定义出发，试图求得监察博弈的混合策略 Nash 均衡。

设代理商偷懒的概率为 x，工作的概率为$(1-x)$。

设委托人检查的概率为 y，不检查的概率为$(1-y)$。

计算代理商的期望盈利函数为

$$\begin{aligned} u_a(\sigma_1,\sigma_2) &= x\{0\cdot y + w(1-y)\} \\ &\quad +(1-x)\{(w-g)y+(w-g)(1-y)\} \\ &= xw(1-y)+(1-x)(w-g) \end{aligned} \tag{2.16}$$

按照 Nash 均衡定义，在给定委托人的混合策略 $\sigma_2=(y,$

$1-y$)条件下，寻求 x 值以使 $u_a(\sigma_1,\sigma_2)$ 达到极大，故在(2.16)中关于 x 求导数并令导数为零，得

$$w(1-y)=w-g \tag{2.17}$$

(2.17)式左端其实就是代理商偷懒时的期望盈利，右端恰为代理商工作的期望盈利。(2.17)式告诉我们，在 Nash 均衡中委托人所取的策略 σ_2（也就是取 y）必须使得代理商在工作或偷懒之间的选择由于平均盈利相等而表现出无所谓态度，故 $y=g/w$。回想在猜谜博弈中，我们求混合策略解的过程也让每个局中人在选取自己的混合策略时使得对方在几个策略之间的选择由于平均盈利一样而采取无所谓态度。可见，在正则型博弈中，求混合策略解与求混合策略 Nash 均衡实质上是一回事。

类似地，通过求委托人的期望盈利并使之达到极大，可以得到 $x=h/w$，它使得委托人在选择检查还是不检查方面持无所谓态度。

综上所述，我们得到监察博弈的混合策略解，或者说得到混合策略 Nash 均衡 $((h/w,1-h/w),(g/w,1-g/w))$。如果监察博弈的研究仅到此为止，那么博弈论在经济问题中的应用就显得有点缺乏活力。现在再进一步，计算得到在达到 Nash 均衡时委托人的期望盈利为

$$v(1-x)-w(1-xy)-hy=v(1-h/w)-w \tag{2.18}$$

它与代理商的生产价值 v、委托人检查费用 h 以及委托人支付给代理商的工资 w 有关，一般地 v 与 h 可视作固定，利用简单的微积分知识可知，若取 $w=\sqrt{hv}$（当然只有在 $\sqrt{hv}>g$ 的情况才有意义）时委托人的平均盈利将达到极大。于是 $\sqrt{hv}$ 将成为委托人与代理商之间签署合同中应付工资的参考值。

2×2 博弈中的划线法很容易推广到两人有限策略空间的博弈中去，只不过在每次划线时要比较多个盈利的大小。现在假定两

个局中人的纯策略空间是一元变量的连续区间，或者纯策略空间虽是有限，但在一段区间内相当多且密集。例如商店的销售额，理论上它基本上是离散的，至少以人民币“分”为最小单位，但是以“分”为单位过分地密集，这类情况可以近似地将纯策略空间处理为连续区间。这类一元变量连续直线的纯策略空间在数学上常称为具有连续统势。如果局中人的纯策略空间具有连续统势，显然“划线法”将无济于事。在这种场合探求 Nash 均衡如何求解是件有意义的工作，因为在社会经济领域内这样的情况比比皆是。譬如两个公司关于某产品的竞争，策略的选择可以是产量，也可以是价格；又譬如我们在海滩占位博弈中，两个局中人关于摊位的设置等等，这些场合的纯策略空间都具有或都几乎具有连续统势。目前我们手头无其他工具，因此还得从 Nash 均衡的定义着手。

假定局中人 1 的纯策略空间为 $x\geqslant0$，局中人 2 的纯策略空间为 $y\geqslant0$，那么平面上第一象限（包括两根正坐标轴）上的每一点都是两人博弈的结局。从局中人 2 的角度分析问题，对于局中人 1 每选择一个策略 x，他作出的最佳反应为策略 y，如果 x 发生变动，那么 y 必然随之有所变动，在平面上将这些点串连起来得到的曲线实质上反映了局中人 2 关于局中人 1 所选策略的最佳反应，我们称这条曲线为局中人 2 的反应曲线，曲线的函数表达（若可能的话）$y=y(x)$ 则称为局中人 2 的反应函数。类似地，可以从局中人 1 的角度同样进行分析，我们得到局中人 1 的反应曲线和反应函数 $x=x(y)$。一般地，这两条曲线在平面上会有交点（也许有时候不止一个交点），Nash 均衡正是“诞生”于这些交点，因为交点 (x,y) 表明两个局中人关于对手所取的策略都作出了最佳反应。

例 2.5 Cournot 竞争模型。

设某产品市场由两家公司垄断，产品的市场竞争中两家公司的策略空间是产量的选择。公司 1 与公司 2 同时从策略可行集合 $Q_i=[0,\infty)(i=1,2)$ 中各选取产量水平 $q_i(i=1,2)$，产品的市场单

价显然与市场上产品的总和密切相关，设为 $p=p(q_1+q_2)$。现以 $c_i(q_i)(i=1,2)$ 表示第 i 家公司生产 q_i 件产品所需成本。于是公司 i 的利润(即盈利函数)为

$$u_i(q_1,q_2)=q_i p(q_1+q_2)-c_i(q_i)\quad(i=1,2)\tag{2.19}$$

记公司 1 与公司 2 的 Cournot 反应函数为

$$r_1:Q_2\rightarrow Q_1\qquad r_2:Q_1\rightarrow Q_2$$

倘若 $u_i(i=1,2)$ 满足一定的理想条件(盈利函数应满足的条件的重要性将在稍后 Nash 均衡存在性问题中涉及)，我们可以求盈利函数的一阶导数并使之等于零(即一阶条件)，从而解得反应函数。例如，$r_2(\cdot)$ 为 Q_1 到 Q_2 的映照，即对于公司 1 选择的产量 q_1，公司 2 的最佳反应函数为 $q_2=r_2(q_1)$，它应满足

$$\frac{\partial u_2(q_1,q_2)}{\partial q_2}=0\qquad(\text{此时将 } q_1 \text{ 视作固定})$$

或者

$$p(q_1+r_2(q_1))+r_2(q_1)p'(q_1+r_2(q_1))-c_2'(r_2(q_1))=0\tag{2.20}$$

同理，$r_1(q_2)$ 应满足

$$p(r_1(q_2)+q_2)+r_1(q_2)p'(r_1(q_2)+q_2)-c_1'(r_1(q_2))=0\tag{2.21}$$

由(2.19)式、(2.20)式解出 $r_2(q_1)$ 与 $r_1(q_2)$，然后求这两条曲线的交点。

现令 $p(q_1+q_2)$ 与 $c_i(q_i)$ 更具体一些，以便进一步理解问题。

令
$$p(q_1+q_2)=\max(0,12-(q_1+q_2))$$

或写作

$$p(q_1+q_2)=\begin{cases}0 & \text{当 } q_1+q_2>12\\ 12-(q_1+q_2) & \text{当 } q_1+q_2\leqslant 12\end{cases}$$

对 $p(q_1+q_2)$ 的上述假设在一定程度上有其合理性，比如总产量 $q=q_1+q_2$ 愈多，将使市场价格减少，产量多到一定程度，产品几乎不值一文。只不过在 p 的定义中，为了演绎上的方便，我们去除了量纲。再令两个公司每生产一件产品的成本为常数 c(c 应当有一

定范畴，这里暂不讨论），即 $c_1(q_1) = cq_1$ 与 $c_2(q_2)=cq_2$。将 $p(q_1+q_2)$与 c_1、c_2 代入(2.19)式、(2.20)式，得

$$r_2(q_1)=(12-q_1-c)/2$$

$$r_1(q_2)=(12-q_2-c)/2$$

Nash 均衡 q_1^*、q_2^* 为这两条曲线的交点，故满足 $q_2^*=r_2(q_1^*)$与 $q_2^*=r_1(q_2^*)$，于是可计算得

$$q_1^*=q_2^*=4-c/3$$

尽管在 Cournot 博弈中，局中人的纯策略空间具有连续统，但它与有限策略空间的囚徒窘境一样存在着个别理性与共同理性的矛盾。就是说，如果两个局中人互相协商合作的话，也许它们的处境会更好一些。为说明这个问题，并使计算简便一些，不妨令 $c=0$，这对整个问题无实质性影响。此时博弈的 Nash 均衡为 $q_1^*=q_2^*=4$，两个局中人的盈利均为 $u_1(4,4)=u_2(4,4)=16$。但若注意到局中人 i 的盈利函数具有形式

$$u_i(q_1,q_2)=q_i(12-q_i-q_2) \qquad (i=1,2)$$

若取 $q_1=q_2=3$，则有 $u_1(3,3)=u_2(3,3)=3\times6=18$。显然如果两个公司精诚合作都限制自己的产量为 3，那么它们将取得更令人满意的盈利 18，它高于取 Nash 均衡(4,4)时的盈利 16。Nash 均衡并没有使双方获得最大盈利。

如果在 Cournot 竞争模型中，两家公司的策略选择不是产量而是价格，那么摆在我们面前的就是 Bertrand 双寡头模型。

例 2.6 Bertrand 双寡头模型。

考虑两个公司生产不同品牌、不同质量与不同包装的同类产品。如果公司 1 与公司 2 分别选择价格 p_1 与 p_2，对于公司 i 的顾客需求量是

$$q_i(p_i,p_j)=a-p_i-bp_j$$

其中 $b>0$ 反映了公司 i 的产品替代公司 j 产品的程度。与 Cournot 模型一样，不考虑固定的生产成本而令边际成本为常数

$c, c<a$。现在两个公司同时选择它们的价格。价格 p_1 与 p_2 的可行集合均取为$[0,\infty)$，因为不可能取负价格。由于不考虑固定生产成本，公司 i 的盈利函数可定义作

$$\begin{aligned}u_i(p_i,p_j)&=q_i(p_i,p_j)[p_i-c]\\&=[a-p_i+bp_j][p_i-c]\qquad(i=1,2)\ (i\neq j)\end{aligned}$$

由于盈利函数光滑可导，仍用求偏导数方法，得

$$\frac{\partial u_i}{\partial p_i}=a-2p_i+c+bp_j=0$$

及

$$p_i=\frac{1}{2}(a+c+bp_j)$$

解联立方程

$$\begin{cases}p_1^*=\dfrac{1}{2}(a+c+bp_2^*)\\[2ex]p_2^*=\dfrac{1}{2}(a+c+bp_1^*)\end{cases}$$

得

$$p_1^*=p_2^*=\frac{a+c}{2-b}$$

$\left(\dfrac{a+c}{2-b},\dfrac{a+c}{2-b}\right)$是该博弈的 Nash 均衡，或称 Bertrand 均衡。显然，我们发现仅当 $b<2$ 时才有意义。

Cournot 均衡与 Bertrand 均衡都是基于盈利函数的可微且严凹特性而从一阶条件导出，有时候遇到的盈利函数没有这样好的性质，但仍存在解，这就需要借助于其他的数学工具了。

§2.6 多重 Nash 均衡

我们已经知道，即使对于 2×2 这样最简单的博弈问题，也可能存在不止一个 Nash 均衡。实际应用工作者也许不喜欢均衡的多重性，因为他们面临着一个棘手问题，既然 Nash 均衡是如何进

行博弈的一个合理预测，那么究竟以哪一个均衡进行预测呢？我们先以一个例子来观察在具体分析中有何尴尬之处。

例 2.7 懦夫博弈(Chicken game)。

两个局中人从一条河的两岸急着同时要过桥，附近没有任何其他的渡河设施。每个人面临着两种选择，要么自己先过桥，要么让对方先过桥。如果双方都采取强硬态度(T)让自己先过，那么走到独木桥中间必然发生冲突(当然我们不考虑他们具有杂技演员本领而安全地交换位置过桥的可能)，其结果将是两败俱伤，假如令这种场合的效用(这里不宜用盈利二字)为-1；如果两个局中人都采取软弱态度(W)，其结果是谁也不过桥，相应的效用自然为零；倘若两个局中人中间有一个强硬另一个软弱，那么强硬的一方立即先过桥取得效用 2，而软弱的一方虽然晚了一点，但毕竟还是过了桥，从而取得效用 1。这些都是两个局中人的共同知识，因此是完全信息静态博弈。其效用矩阵如图 2.16 所示。

		局中人 2	
		T	W
局中人 1	T	$-1,-1$	$2,1$
	W	$1,2$	$0,0$

图 2.16

利用划线法立即可知，结局(T,W)与(W,T)都是博弈的纯策略 Nash 均衡。不难验证$((1/2,1/2),(1/2,1/2))$是懦夫博弈的混合策略 Nash 均衡。一个看上去很简单的 2×2 博弈模型，居然有三个所谓的合理预测，的确使人有点拿不准主意。暂且抛开混合策略不谈，在(T,W)与(W,T)两个纯策略之间究竟相信哪一个更为合理一些呢？倘若我们有些额外信息，譬如局中人 1 一向比较专横，从不愿意让人，而局中人 2 性格属于胆小谨慎且对人比较谦让的一类，如果双方也很了解对方的性格与历来的处世习惯，那么我

们将毫不犹豫地预测结局(T,W)将会发生。但是如果我们对两个局中人一无所知，他们相互之间也不认识、不了解，在此之前更从未玩过类似的"懦夫博弈"游戏，要准确地预测结局自然是件很难的事情。我们没有明确的方法可以用来协调局中人的期望与行动，此时，博弈最终出现的不管是(T,W)也好，还是(W,T)也好，人们都不会表现出什么惊异。相反，在此种场合，倘若有人断言博弈的结局应当是(T,W)，反倒会使人感到吃惊，或者认为此人过于武断，或者嘀咕"他一定在事前获得过什么其他的信息"。

Nash 均衡多重性令人难堪之处就这样明明白白地摆在我们面前。从对懦夫博弈的分析中可以看到，在博弈之前对局中人的有关信息的收集似乎是个关键。然而，世界上的事情并非可以想当然，日常生活中的博弈极少有可能在事前收集到需要的信息。因此从多重 Nash 均衡中挑选一个作为合理且正确预测的一般性规律几乎无法找到。但是一些博弈论专家在某些特殊场合所作的建议或者提出的若干准则对人们有所启发。我们将在下面略作介绍。

1. 聚焦

Schelling 于 1960 年提出的"焦点"理论指出在某些"日常生活"情况，局中人可以通过利用由策略形式提供的信息而协调在特殊的均衡上。例如，策略的名字可能有某些共同理解的"聚焦"能力。假定玩一个游戏，要求两个局中人独立地写出$(-1/2,1/2)$中的任意一个实数，倘若他们两人写出的实数相吻合，则每人奖励 100 元，如果不合，则每人被罚 10 元。显然每个局中人的策略空间是$(-1/2,1/2)$，对于局中人 1 的任意一个选择$x\in(-1/2,1/2)$，局中人 2 的最佳反应应该是$y=x$，反之，对于局中人 2 的任意选择$y\in(-1/2,1/2)$，局中人 1 的最佳反应也应该是$x=y$。因此，对于一切$t\in(-1/2,1/2)$，(t,t)构成了博弈的 Nash 均衡，也就是说，博弈有无穷多个 Nash 均衡（且 Nash 均衡集合具有连续统势），从表面上看，似乎赢钱的可能性相当多，但是从概率论角度

看，只要两个局中人从$(-1/2,1/2)$中随机地取数的话，(x,y) $(x\neq y)$的发生概率远远大于(t,t)发生的概率(其实前者为1而后者为0)。博弈中的局中人是有理性的，为了想赢钱，他们不会采取这种"随机"策略，因为任意选取一数几乎没有赢钱的希望，比如某人取$x=0.0032$，根本不可能设想对方能预测到自己会取这个数，然而策略形式"0"却是个可能协调双方期望的焦点，也许是双方都容易取的数，所谓"心有灵犀一点通"，$(0,0)$是个较合理的预测结局。类似的例子可以举出很多，两个人分一块蛋糕，每人独立地提出自己想分得的百分比，如果两个百分数相加正好等于1，那么各人可分得自己所要的份额，否则就没有份。这里，$(0.5,0.5)$就是个焦点。其实，这个焦点很容易得到，每个局中人为极大化盈利起见，提出的蛋糕要求总是大于等于$1/2$，因此$1/2$就是一个双方都能接受也最可能想到的聚焦点。

以聚焦形式建议的预测，其根本原因在于各种各样策略的"焦点"依赖于局中人的文化背景、习惯或历史经验。

2. 风险占优

从多重Nash均衡中选取一个合理的预测常常依赖于预测风险的大小。众所周知，预测毕竟不一定是博弈的最终结果，纵然我们使用的是Nash均衡，总归存在风险。人们通常比较倾向于接受预测风险较小的结局，其理由是不言而喻的。

例如，在§1.2共同投资问题中，倘若两家公司改为$n(>2)$家公司。这个博弈至少有两个纯策略Nash均衡：

(1)所有n家公司都投资大工程

对于公司$i(i=1,2,\cdots,n)$来说，在其余$(n-1)$家公司都投资大工程时，它的最佳反应毫无疑问地是选择投资大工程，因为一旦它不这样做，除了其余$(n-1)$家公司将受到重创之外，它本身从投资小工程收得的回报也将大大低于它投资大工程时应有的回报。也就是说，公司i只有与大家一起坚持投资大工程，才能极大化自

己的盈利或效用，它不会主动偏离这个策略剖面。由此，n 家公司全部选择投资大工程是 Nash 均衡。

(2)所有 n 家公司都投资小工程

任何一家公司都会考虑到，只要在其余的($n-1$)家公司至少有一家转而投资小工程的话，它的最佳反应一定是投资小工程，否则它的资金将被套住。n 家公司都是理性的，因此"它们都投资小工程"也是 Nash 均衡。

这两个纯策略 Nash 均衡中哪一个作为博弈的预测更合理呢？全部投资大工程，收效甚大而达到皆大欢喜，大家投资小工程收效则少得多，也许仅够"小康"。但是在收效上的差异不能作为选取哪个 Nash 均衡作为预测结局的依据，因为在收效大的均衡中各公司承担的风险也大，一旦有个别局中人发生"背叛"，或者在作出决策问题时出了点小小的差错，其余局中人连小收益也得不到。那么都投资大工程会造成多大的风险呢？或者问，其可能性有多大呢？假定如下所知为共同知识：各公司在独立考虑问题时通常以 1/2 或更多的概率选择投资大工程。倘若 $n=2$，即 § 1.2 中情况，两家公司都投资大工程的概率应为 1/4 或更大一些，而两家公司都投资小工程的概率为弱于 1/4。从预测最终结局来说，显然"都投资大工程"优于"都投资小工程"。对于每一个局中人来说，由于其对手投资大工程的可能性大于 1/2，因此他预测对手可能投资大工程，他自己的最优策略也就因此确定为投资大工程。但是当 n 发生变化时，概率的计算也随之变化。例如，当 $n=17$ 时，结果就完全相反。此时，每一家公司为了尽可能少冒一些风险，只有当它认定"其余 16 家公司都投资大工程"这件事至少具有 1/2 以上概率时，它才会将投资大工程作为自己的最佳选择。设每家公司独立思考时选择投资大工程的可能性为 $p(0<p<1)$，那么这个"前提"要求 $p^{16}\geqslant 1/2$，即要求 $p\geqslant 0.96$，这是一个相当大的概率。分析蕴含了这样一个意思，要求每家公司几乎肯定投资大工程，此时投资大

工程才是每家公司的最佳选择。由此看来，n 较大时，预测“全部投资大工程”作为博弈的结果显然具有很大的风险。用 Harsanyi 与 Selten（1988）的话来叙述，从风险角度来看，“全部投资小工程”优于“全部投资大工程”。

3. Pareto 最优均衡

共同投资博弈中，所有公司全部投资大工程，从盈利收益方面，对任何一个局中人无疑是最优的结局。这类结局的特点是，在不损害他人的前提下，局中人将不可能再增加自己的利益，这是一个有效结局，也是数理经济学所追求的 Pareto 最优。按理讲这是一个人人喜欢的结局，但是，Harsanyi 与 Selten 的风险占优恰好提出了博弈并不总是以 Pareto 最优均衡结局作为其合理预测。现在设想如果博弈的确存在着 Pareto 最优均衡，又假定局中人在博弈之前互相交谈，那么他们很可能在 Pareto 最优均衡方面进行协调。这种交谈不负任何法律责任，因为毕竟仅仅是交谈而不是订合同或契约。不过这类交谈有可能确保局中人在运用 Pareto 最优均衡策略时只冒较低风险。譬如对“共同投资大工程的概率 p 可能有所提高”一事抱有信心，导致“全部投资大工程”的可能性增大，但无绝对把握。

例 2.8 图 2.17 所示博弈。

局中人 2

局中人 1		L	R
	U	9,9	0,8
	D	8,0	7,7

图 2.17

划线法显示，博弈存在两个纯策略均衡：(U,L)与(D,R)。可计算出混合策略 Nash 均衡的平均盈利比 7 与 9 更低。毫无疑问，结局(U,L)Pareto 优于其他均衡策略。那么(U,L)是否为如何进

行博弈的最合理的预测结果呢?如果在博弈之前,两个局中人不发生任何交流,双方关于对手只有“他是个理性人”这样单纯的信息。由于盈利矩阵完全已知,既然预测结局从(U,L)与(D,R)中选取一个(那个平均盈利差的混合策略自然不在考虑之列),由于大家都希望自己的盈利达到最大,(U,L)的盈利使得它有可能成为聚焦点。

现在从风险占优的角度来考虑,按局中人1的观点,策略D比策略U更“安全”一些,因为局中人1只要选取了策略D,不管局中人2如何行动,局中人1至少可以获得盈利7,或者更好一些(8)。但若他取策略U,尽管他可能获得博弈的最高盈利9,然而也存在着落得一无所有的可能(即当局中人2取R时)。那么局中人2取R的可能性究竟有多大呢? 不妨令此概率为y,计算局中人1取U时的期望盈利得

$$9(1-y)+0\cdot y=9-9y$$

而他取策略D时的期望盈利为

$$8(1-y)+7y=8-y$$

当$y>1/8$时,有$9-9y<8-y$,这表明,如果局中人1预测到局中人2取策略R的概率大于1/8的话,从平均盈利考虑,局中人1应采取策略D。注意到盈利矩阵关于两个局中人是对称的,同样的讨论告诉我们,如果局中人2预测局中人1取D的可能性大于1/8的话,局中人2应当采取策略R以获得较高的期望盈利。1/8是个较小的数,因此一般来说,从风险占优角度,(D,R)优于(U,L)。

对例2.8的讨论揭示了一个有趣的事实,聚焦与风险占优得到了不同的Nash均衡作为博弈的预测。看来,在多重Nash均衡情况下,预测博弈究竟会发生什么样的结果,的确是件不确定的事。

转而我们考虑在博弈之前局中人预先进行交流的情况。如果局中人1流露愿采取策略U的想法,显然这种想法有益于局中人

2，因为他要么得 9 要么得 8。而他取 L 将使自己获最高盈利 9，利用交谈，局中人 2 将告诉局中人 1 有关自己的理性想法，从而坚定局中人 1 取策略 U 的决心，这使得 Pareto 最优均衡(U,L)的可能性大大增加。不过在交流之后每个局中人又怎么相信对方的保证呢？假如没有协议（尤其是具有法律效用的协议）的约束，极有可能最终的结局是另一个 Nash 均衡(D,R)。就是说，局中人的预先交流会增加 Pareto 最优均衡（若存在的话）的可能性，但并非使人非得相信一定会出现 Pareto 最优均衡。

4. 防联盟(coalition-proof)均衡

以 Pareto 最优均衡作为自然预测的想法遇到另外一个困难在多于两个局中人的博弈中体现出来。这是因为三个及三个以上的局中人，就有可能部分人结成"联盟"，在极大化联盟成员利益的同时损害了其他局中人的利益，这样的结局显然有悖于 Pareto 最优均衡的基本精神。

先以一个简单的例子逐步展开讨论。

例 2.9 考虑甲、乙、丙三人博弈，盈利矩阵如图 2.18 所示（每格的三个数从左到右依次表示甲、乙、丙在该局时的盈利）。

丙取 A：

甲 \ 乙	L	R
U	0,0,10	−5,−5,0
D	−5,−5,0	1,1,−5

丙取 B：

甲 \ 乙	L	R
U	−2,−2,0	−5,−5,0
D	−5,−5,0	−1,−1,5

图 2.18

尽管本例有三个局中人和两个盈利矩阵，仍然可以通过划线法来求得纯策略均衡。例如，在同时固定乙取 L 和丙取 A 的条件下，比较甲取 U 与 D 的盈利大小，显然 $0>-5$，则在 0 下面划一

短线。又例如，在固定甲取U且乙取L的条件下，丙取A的盈利为10，而取策略B时得盈利0，于是我们就在10下划一短线，如此等等。我们最后去找那些三个元素均划有短线的结局，易知(U,L,A)与(D,R,B)是纯策略Nash均衡。本例题的原意是讨论(U,L,A)与(D,R,B)中哪一个比较合理，因而暂不涉及混合策略Nash均衡，但对这样的例题求解混合策略也许有利于对混合策略解的进一步理解。

令x表示甲取U的概率，y表示乙取L的概率，z表示丙取A的概率。

甲取U时的期望盈利为

$$0-5(1-y)z-2y(1-z)-5(1-y)(1-z)$$

甲取D时的期望盈利为

$$-5yz+(1-y)z-5y(1-z)-(1-y)(1-z)$$

所谓混合策略均衡，意即乙、丙分别取y、z，应当使甲在取U与取D这两者之间感到无所谓，即上述两个期望盈利必须相等，得

$$4yz-2z+7y-4=0 \tag{2.22}$$

同理，考虑乙的期望盈利可得

$$4xz-2z+7x-4=0 \tag{2.23}$$

考虑丙的期望盈利可得

$$xy=(1-x)(1-y) \quad \text{即}\ 1-x-y=0 \tag{2.24}$$

综合(2.22)式、(2.23)式与(2.24)式，不难解得$x=y=z=1/2$。因此$((1/2,1/2),(1/2,1/2),(1/2,1/2))$是混合策略Nash均衡，相应的期望盈利为$(-3,-3,5/4)$。

现在回到纯策略Nash均衡，显然(U,L,A)Pareto优于(D,R,B)(尤其它也优于混合策略均衡)。那么(U,L,A)是否可成为甲、乙、丙三个局中人明显的聚焦点呢？不妨设想(U,L,A)就是博弈的预测解，此时我们使局中人丙的选择固定于策略A，于是甲乙

二人之间的条件博弈的盈利矩阵如图 2.19 所示。

乙

		L	R
甲	U	0,0	−5,−5
	D	−5,−5	1,1

图 2.19

该条件博弈的 Pareto 最优均衡是(D,R)。于是甲乙两人期望局中人丙采取策略 A,在这基础上甲与乙协调二人的行动于条件博弈的 Pareto 最优均衡。事实上,只要有可能,他们的确应当这样协调自己的行动,其结果却推翻了原博弈的一个均衡(U,L,A)。

发生这样的情况相当于局中人甲与乙组成了一个联盟,在与丙的博弈中,联盟一方互相协调尽可能地极大化联盟各个成员的盈利,作为联盟的对手,局中人丙是理性的,他清楚地知道如果博弈三方在博弈过程中互相独立地行动,那么(U,L,A)与(D,R,B)是两个可能的"好"结局,相比较(U,L,A)也许对三方都更有利一些。但是如果根据这样的分析,他选择了策略 A,有可能引起甲乙二人结成联盟(由于甲乙为了自身利益的极大化,这种可能性非但存在而且并不小),从而预测结果看来不会是(U,L,A)。逻辑分析表明,甲乙协调于(D,R)的可能极大,于是丙为了保障自己的利益,将毫不犹豫地转向选择策略 B。假如丙选择 B,此时甲乙二人的条件博弈的 Pareto 最优均衡仍为(D,R),于是博弈的合理预测看来应当是 Nash 均衡(D,R,B),它不是 Pareto 最优均衡。(D,R,B)有效地防止了甲乙二人可能的联盟,从而避免了丙的损失。因此我们说,在防联盟均衡这一层意义上,(D,R,B)优于(U,L,A)。

或许有些读者会指出,倘若我们改换考虑的角度,先固定甲的

策略选择为 U，而由乙与丙互结联盟，情况可能未必如前。事情的确是这样，如此考虑推断的话，(U,L,A)是合理预测。也就是说，从甲的角度，他似乎未必要防止乙与丙结盟，因为当甲固定 U 时，不管乙与丙是否结盟，他们的“条件”Nash 均衡仍为(L,A)，因而不违背原博弈的 Nash 均衡(U,L,A)。

对博弈的预测需要从整体出发，因此防联盟均衡也应全面考虑。像本例那样的三人有限博弈，究竟选择哪一个 Nash 均衡作为合理预测呢?从防联盟均衡意义上看，应当在面定任何一个局中人的策略选择时，其他两个局中人将协调在条件博弈的 Pareto 最优均衡上，如果这样协调的结果偏离了原始的 Nash 均衡，那么这个 Nash 均衡(不管它是否是三人博弈的 Pareto 最优均衡)就不能成为合理的预测，例如(U,L,A)。假如协调并没有背离原来的 Nash 均衡，那么这个结局可能成为一个合理的预测，本例中的(D,R,B)就是这样的实例。概括一句话，我们要求三人博弈中任何两人联盟在他们的 Pareto 最优均衡(倘无，则在 Nash 均衡)上协调并不背离原先博弈的 Nash 均衡，这样的 Nash 均衡就是防联盟均衡，它可作为博弈的一个合理预测。

上述思想可以推广到多于三人的博弈问题：n 个局中人的博弈，其可能的合理预测，首先要求所预测的结局是 Nash 均衡，在那里每一个局中人都在其余$(n-1)$个局中人策略选定的基础上由于已经极大化自己的盈利而不愿偏离。从“防联盟”的观点叙述，即没有任何“一人联盟”会偏离 Nash 均衡策略。进一步固定任意$(n-2)$个局中人的策略，观察余下的局中人在给定的条件博弈中协调的 Pareto 最优均衡(倘无最优，则为 Nash 均衡)是否偏离原博弈 Nash 均衡的纯策略剖面。即原始博弈的一个 Nash 均衡如果可能成为一个合理预测，我们要求不存在任何一个两人联盟愿主动偏离它；类似地，我们要求任何一个三人联盟不偏离原博弈的 Nash 均衡……一直下去，直至任何一个$(n-1)$人联盟不会偏离原

博弈的 Nash 均衡。总而言之，在多人博弈中，如果存在多重纯策略 Nash 均衡，哪一个可以作为合理预测呢？从防联盟观点出发，任何 $k(1\leqslant k\leqslant n-1)$人联盟都不会发生背离现象的 Nash 均衡是一个合理预测，符合这种推理的预测结局称作防联盟均衡，这种想法是由 Bernheim、Peleg、Whinston 于 1987 年提出的。

§2.7 相关均衡

我们已经引进了 Nash 均衡这个在博弈论中具有极其重要的意义与价值的概念，通常它在下述范围才有意义：局中人独立地选择自已的策略。事实上在不少博弈中有时候局中人的策略选择是相关的，例如，可与一个“信号设置”有关。

假如古代两位势均力敌的将军大战数百回合，每个人气喘吁吁，谁都知道胜不了对手，他们各面临着两个可供选择的策略：继续打下去或立即停止。双方停止也许是个有效结局，但在没有任何信号的情况下，预测结局是如下 Nash 均衡：双方继续打，任何一方主动后撤将表示退却，因此没有人会主动偏离“继续打”的局面。倘若一方或双方鸣金收兵，这是一个双方都可接受的(体面的)信号，二人将会停止战斗。这里，最终的策略与信号有关，也就是说局中人的策略选择是相关的，但双方都由此获得了好处。

两家公司进行市场竞争，假定出现某种信号，例如亚洲金融风暴，在双方观察到这个信号之后所进行的策略选择是相关的。预测的结局也将与两家公司独立地选择策略时可能有所不同。一般地，通过“相关装置”，使局中人获得更多的信息，常常使他们从中获益。

本小节将对相关均衡这一概念进行初步探讨。先考虑由 Aummann 提出的例子(见图 2.20)。

乙

		L	R
甲	U	5,1	0,0
	D	4,4	1,5

图 2.20

容易验证，在完全信息下，如果甲乙同时（独立地）行动，此博弈有三个均衡：(U,L)，(D,R)与混合策略$((1/2,1/2),(1/2,1/2))$，它们的盈利向量分别为$(5,1)$，$(1,5)$与$(2.5,2.5)$。如果他们在博弈之前建立一个信号装置，比如可以共同观察到"抛一枚公正的钱币"的结果。假定甲乙双方约定：若硬币呈正面，甲取U而乙取L；若硬币呈反面，则甲取D而乙取R。这种约定相当于他们在两个纯策略 Nash 均衡之间随机化地选择，由于钱币是公正的，呈正反面的概率各为 1/2，于是可以算得这个相关装置使甲乙双方的期望盈利均为

$$1/2\times 5+1/2\times 1=3$$

盈利$(3,3)$显然大于混合策略均衡的期望盈利$(2.5,2.5)$，现在将相关装置稍稍扩大到较一般情况。设所掷钱币是不均匀的，呈正面的概率为$\lambda(0<\lambda<1)$，呈反面的概率为$(1-\lambda)$。双方的约定仍如前，那么甲的期望盈利为

$$\lambda\cdot 5+(1-\lambda)$$

而乙的期望盈利为

$$\lambda+5(1-\lambda)$$

它们构成的期望盈利向量实际上可以表示成

$$\lambda\times(U,L)\text{的盈利向量}+(1-\lambda)\times(D,R)\text{的盈利向量} \quad (2.25)$$

当$\lambda\in[0,1]$变化时，(2.25)式的全体构成两个纯策略 Nash 均衡盈利向量的凸包。显然，抛硬币这样的信号装置可以使局中人获得该凸包中任意期望盈利向量，但不可能获得凸包之外的任何盈利

向量。这句总结性的语言可以推广到更一般的情况，因为共同可观察的随机现象不只是抛硬币。如果将"抛硬币这样的相关装置"改为"通过使用共同可观察的随机变量"，上述结论当然仍然成立。

相关装置使局中人获得更多信息从而获得更好的(期望)效用，倘若将信号装置设计得更巧妙一些，事情也许会更好一些。例如图 2.19 所示的博弈中，假设甲乙分隔两处，事前知道由丙从 A、B、C 三张卡片中等可能地抽取一张，若为 A 则出示给甲看，若为 C 则出示给乙看，若抽得 B 则不出示给任何一个局中人。因此甲在看不到牌时就认为可能发生了"B"或者"C"，同样乙在看不到牌时将会认为装置不是发生"A"就是发生"B"。

在上述相关装置的帮助下，该博弈的 Nash 均衡为：

局中人甲被告知 A 时取策略 U；

局中人甲被告知(B,C)时取策略 D；

局中人乙被告知 C 时取策略 R；

局中人乙被告知(A,B)时取策略 L。

首先要做的事情是确认在上述行动中局中人不会发生违背与偏离。事实上，当甲看到 A 牌时，根据事前约定，他知道乙一定认为发生了(A,B)，于是乙取策略 L，此时甲的最佳反应必然是 U。反之，如果甲没有见到牌，他认为(B,C)发生。由于甲无法区别究竟发生 B 还是 C，所以他应该想到此时乙可能面临两种状态，要么他看到了 C 牌，要么由于事实上发生的是 B 牌，因而乙也没有看到任何牌，于是乙认为发生的是(A,B)，抽得 A、B 与 C 牌的等可能性使得甲相信乙所认为的这两种状态是等可能的，因而甲自然地期望乙以等概率取策略 R 与 L。计算表明此时甲取 U 与 D 都可获得期望盈利 2.5，但比较而言，取策略 D 不会发生结果为一无所有的极端局面，故 D 比 U 更稳妥一些，它可能比 U 更容易被甲所接受，也就是说 D 是甲在此时的最优反应。同样的分析显然也适用于乙，因此局中人甲与乙均不会偏离这个结局。

归纳一下，我们发现在这个均衡解中：

若抽得 A 牌，局中人甲取 U，局中人乙取策略 L。

若抽得 C 牌，局中人乙必取策略 R，而此时局中人甲因认为可能发生 (B,C) 而选取策略 D。

若抽得 B 牌，由于甲与乙均观察不到牌，甲认为发生了 (B,C) 而乐意取 D；而乙认为发生了 (A,B)，同样的分析可知乙乐意取 L。

由于 A、B、C 三种状态是可能的，因此我们建立了局中人的选择为相关的均衡——相关均衡：

各以 1/3 概率取 (U,L)、(D,L) 和 (D,R)。

注意：这个相关均衡避免了“坏”结局 (U,R) 的发生。在这个新的相关均衡中，局中人甲与乙的平均盈利分别为：

甲：$1/3\times5+1/3\times4+1/3\times1=10/3$

乙：$1/3\times1+1/3\times4+1/3\times5=10/3$

盈利向量 $(10/3,10/3)$ 显然优于 $(3,3)$，而且它不属于原博弈 Nash 均衡盈利向量的凸包。

我们可以看到一个有趣的事实，建立了额外的信号装置不排除博弈原来的均衡 (U,L) 与 (D,R)，这是因为信号并不影响博弈的盈利矩阵。

在相关均衡问题中，局中人限制了自己的信息，当对手获知他这样做的话，就可以被诱导而以合理的形式进行博弈，从而使得局中人获得好处。下面的相关均衡例子阐述了这一点。

例 2.10　三人博弈的盈利矩阵（见图 2.21）。

划线法告诉我们，该博弈唯一的纯策略 Nash 均衡是 (D,L,A)，相应的盈利向量为 $(1,1,1)$。现在试图建立一个信号装置以使局中人相关地选择自己的策略从而提高期望盈利。

乙

甲	L	R
U	0,1,3	0,0,0
D	1,1,1	1,0,0

A

乙

甲	L	R
U	2,2,2	0,0,0
D	2,2,0	2,2,2

B

乙

甲	L	R
U	0,1,0	0,0,0
D	1,1,0	1,0,3

C

丙

图 2.21

设想建立一个抛均匀钱币的信号装置，如果局中人甲与乙可以观察到钱币呈现正面(H)还是呈现反面(T)，而丙对信号什么都看不到。在该信号装置下博弈的 Nash 均衡为：

甲：若钱币呈 H，取 U；若钱币呈 T，取 D。

乙：若钱币呈 H，取 L；若钱币呈 T、取 R。

丙：不管信号怎样均取 B。

根据上述约定，局中人丙面对着甲、乙二人的行动分布：以1/2概率甲乙取(U,L)且以1/2概率甲乙取(D,R)。此时丙取 A、取 B 与取 C 的期望盈利分别是 3/2、2 和 3/2，因此丙的最佳反应自然是 B。在本相关均衡中注意到丙看不到信号的重要性，否则当丙看到钱币呈正面时，他知道甲、乙取(U,L)，那么他将违背约定而取策略 A 使自已的盈利增加至 3。同样地丙若看到钱币呈反面，他将会取策略 C。这等于说，信号装置倘稍有变动，那么相应的相关均衡将随之发生变化。

我们已经介绍了两个相关均衡的例子，得到的结论是，相关均衡实际上相当于在适当巧妙的信号装置下博弈的“条件”Nash 均衡。

第三章　Nash 均衡存在性定理

第二章中我们介绍了博弈论里极其重要的概念——Nash 均衡，对于将 Nash 均衡作为进行博弈的合理预测有了基本认识，也了解了如何求解纯策略与混合策略 Nash 均衡，在多重 Nash 均衡存在的时候如何考虑寻找更为合理的结局作为预测，在可能的情况下巧妙的信号装置的建立又怎样地提高局中人的期望盈利等等。这些内容固然很重要，但是更重要的是需要知道在正则型（或策略型）博弈中是否存在 Nash 均衡。“皮之不存，毛将焉附”，不解决存在性问题，则可能使其余一切问题都失去意义。本章的目的在于证明 Nash 均衡存在性定理。

我们已经知道，博弈未必存在纯策略 Nash 均衡，因此存在性问题通常在混合策略范畴内讨论。定理的证明将涉及一些数学技巧，对于那些只希望关心博弈论在社会经济领域中的应用的读者，可以跳过本章所使用的证明而继续阅读其他内容。不过，本章所使用的证明方法原则上只用到较初等的数学，只要稍具数学知识的读者基本上都可以读懂。因此建议读者尽可能地阅读本章，这样可以对 Nash 均衡有更深入的了解。我们从最通俗的叙述入手。

§3.1 Cournot 竞争的 Nash 均衡可视作市场调整的结果

回忆 Cournot 竞争模型的求解，我们试图求出两条反应曲线

或反应函数 $r_1(q_2)$ 与 $r_2(q_1)$，假如它们存在交点，则交点坐标相应的结局就是所需寻求的 Nash 均衡。这种方法是从定义出发，因此在理论上无懈可击。可是在具体操作中，公司 i 不可能知道公司 j 的产量选择，它们通过在市场竞争中所了解的对手的实际产量不断地调整自己的策略。例如，局中人 1(或 2)在刚开始时(不妨记为 0 时期)确定自己的初始产量 q_1^0(或 q_2^0)，局中人 2(或 1)观察到这个 q_1^0(或 q_2^0) 后必定在下一个时期作出自己的最佳反应而选定 $q_2^1=r_2(q_1^0)$(或 $q_1^1=r_1(q_2^0)$)，$r_2(\cdot)$(或 $r_1(\cdot)$)即为 Cournot 反应函数。局中人 1(或 2)在时期 1 看到 q_2^1(或 q_1^1)，也会在时期 2 相应地作出最佳反应 $q_1^2=r_1(q_2^1)=r_1(r_2(q_1^0))$，$q_2^2=r_2(q_1^1)=r_2(r_1(q_2^0))$……可以设想这个过程将一直延续下去，双方不断地依据对手的产量进行策略调整，市场调整到一定的时候，两个局中人的产量将“稳定”在某一状态(q_1^*,q_2^*)，也就是说，根据 q_1^*，局中人 2 的下一时期最佳反应为 q_2^*，而对于 q_2^*，局中人 1 在下一时期的最佳策略选择是 q_1^*，这里的(q_1^*,q_2^*)就是 Cournot 竞争的 Nash 均衡。用数学上较严格一些的术语，在市场调整过程中，两个局中人各有一系列产量策略选择：

$$q_1^0,q_1^1,q_1^2,\cdots,q_1^n,\cdots$$

$$q_2^0,q_2^1,q_2^2,\cdots,q_2^n,\cdots$$

如果这两个序列各收敛于 q_1^* 与 q_2^*，显然应有

$$q_1^*=r_1(q_2^*),q_2^*=r_2(q_1^*) \tag{3.1}$$

市场调整达到或者几乎达到这个程度，那么市场处于一个相对稳定的阶段，作为任何一方局中人，谁也不愿主动偏离这个状态以使自己蒙受损失。这就是我们所关心的均衡。

两个局中人在时期 t 通过选择一个关于其对手在$(t-1)$时期行动的最佳反应从而调整各自在市场竞争中的策略。可以用向量形式来表示这一动态过程：

$$q^t=(q_1^t,q_2^t)=(r_1(q_2^{t-1}),r_2(q_1^{t-1}))=f(q^{t-1}) \tag{3.2}$$

f 是将$(t-1)$时期的策略向量 q^{t-1}映射到 t 时期的策略向量 q^t，因此 f 实际上是二元向量到自身的一个映射。当 t 越来越大时，倘若 q_1^t、q_2^t 收敛的话，也就是向量 q^* 收敛，取极限应有

$$q^*=f(q^*) \tag{3.3}$$

(3.3)式表明，q^* 实际上是映射 f 的不动点。如果是 n 个局中人之间的博弈，(3.2)式成为

$$\begin{aligned} q_1^t &= (q_1^t, q_2^t, \cdots, q_n^t) \\ &= (r_1(q_1^{t-1}, q_2^{t-1}, \cdots, q_n^{t-1}), \cdots, r_n(q_1^{t-1}, q_2^{t-1}, \cdots, q_{n-1}^{t-1})) \\ &= f(q^{t-1}) \end{aligned} \tag{3.4}$$

此时 f 为 n 维向量空间到自身的映射，倘若 f 存在不动点，即存在策略向量 q^*，使得

$$q^*=f(q^*)$$

这个不动点 q^* 其实就是 n 人博弈的 Nash 均衡。

由此可见，博弈的 Nash 均衡存在与否完全取决于映射 f 是否存在不动点，因此这与 f 的性质有着密切的关系。映射 f 是由策略的形式以及盈利函数所确定的，这是因为 f 与诸反应函数 r_i $(i=1,2,\cdots,n)$有着极其密切的关系。对诸 r_i 及映射的区域等有什么样的要求才能使映射具有不动点呢？解决此类问题的一个有力工具是 Brouwer 不动点定理，这是一个人们非常容易理解但是一般人较难作出证明的著名数学定理。

§3.2 Brouwer 不动点定理

先以一个简单且通俗的例子来说明 Brouwer 不动点定理，一杯牛奶静止地放在桌上，慢慢地并且连续地搅动和旋转杯中的牛奶，然后让牛奶逐渐地自行静止下来。Brouwer 指出，在这杯重新处于静止状态的牛奶中至少有一个点恢复到尚未搅动前它在杯中原先的位置。

用正规的数学语言来描述 Brouwer 不动点定理：

如果有一个有界闭凸集合被连续地映射到自身，那么至少有一个点被映射到自身。

Brouwer 不动点定理的证明对于非数学专业的人是非常困难的，假如读者具有一点拓扑学的有关知识，对问题的展开将有所帮助。我们在这里给出的是初等证明。

首先介绍单纯形概念，单纯形是如下点的集合：

$$\begin{cases} x = x_0v^0 + x_1v^1 + \cdots + x_Nv^N & (3.5) \\ x_0 \geqslant 0, x_1 \geqslant 0, \cdots, x_N \geqslant 0, x_0 + x_1 + \cdots x_N = 1 & (3.6) \end{cases}$$

$v^0, v^1, \cdots, v^N$ 实际上是该单纯形的顶点。称 $x_0, x_1, \cdots, x_N$ 为 x 的重心坐标。为了直观地理解单纯形，不妨从较小的 N 开始。$N=1$，此时单纯形是一线段，顶点为线段的两个端点，x 为线段上任意一点。事实上，线段上的任意 x 可以表示为两个端点的凸线性组合。如果 $N=2$，顶点有三个，我们的单纯形集合成为平面上的点，三个顶点所构成的三角形的边与内部的所有点显然可以用(3.5)式、(3.6)式来表示，因此 $N=2$ 的单纯形实质上是平面上的三角形。不难想象，$N=3$ 实际上是通常三维空间中的四面体。

我们先在单纯形上证明 Brouwer 不动点定理。

定理 3.1 如果 $f(x)$ 连续地将一个非退化的单纯形映射到自身，则至少存在一个不动点 $x^* = f(x^*)$。

如果用重心坐标来描述，令 $x^* = x_1^* v^0 + \cdots + x_N^* v^N$，又记 $f_k(x)$ 表示点 x 的映象 $f(x)$ 相应于 v^k 的重心坐标（$k=0,1,\cdots,N$），那么该点应满足

$$x_k^* = f_k(x^*) \quad (k=0,1,\cdots,N) \tag{3.7}$$

定理的证明无非说明至少有一点 x^* 满足(3.7)式，其实(3.7)式有一个等价的关系式：

$$x_k^* \geqslant f_k(x^*) \quad (k=0,1,\cdots,N) \tag{3.8}$$

我们在以下的证明中主要是寻找满足(3.8)式的点，因此必须先证明(3.7)式与(3.8)式之间的等价性，等价性完全是因为 x_k^* ($k=0,1,\cdots,N$)是重心坐标而产生的。(3.7)式蕴含着(3.8)式是不言而喻的，因为"大于等于"当然包含了"等于"这一层意思。现在只要说明(3.8)式也蕴含了(3.7)式，自然明白了它们的等价性。假如点 x^* 满足(3.8)式，由于重心坐标的特点，应有

$$\sum_{k=0}^{N} x_k^* = 1 = \sum_{k=0}^{N} f_k(x^*) \tag{3.9}$$

又因为对一切 $k(=0,1,\cdots,N)$有 $x_k^* \geqslant 0$，$f_k(x^*) \geqslant 0$，(3.8)式与(3.9)式的同时成立表示只能有 $x_k^* = f_k(x^*)$，$k=0,1,\cdots,N$，即(3.7)式成立。

现在考虑在单纯形的每个边界面上所有的点，都不可能发生某个不等式 $x_k > f_k(x)$ 的情况。首先注意到，对于每一个顶点 v^k 而言，它的重心坐标只有在相应于 v^k 前的系数等于 1，即 $x_k=1$，相应于其余 $v^j(j\neq k)$的重心坐标等于 0。而与 v^k 完全相对面的边界上的点其相应于 v^k 的重心坐标 x_k 必定等于 0。图 3.1 对 $N=2$ 的情况进行了描述。

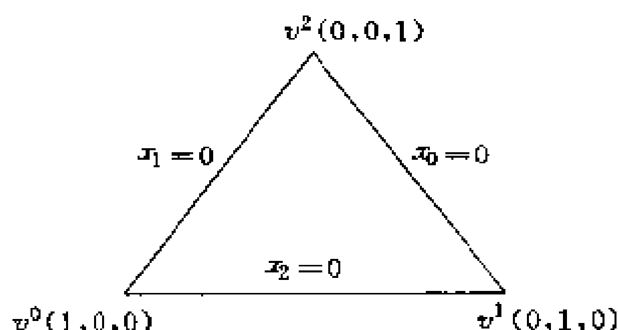


图3.1

设 x 为非不动点，具有映象 y，即 $y=f(x)$，由于 $x\neq y$，因此 x 与 y 的重心坐标不全相等，或者说存在某些重心坐标 $x_i\neq y_i$，由于重心坐标必须满足

$$\sum_{k=0}^{N} y_k = \sum_{k=0}^{N} x_k = 1$$

自然存在某些 $x_i>y_i$，而另一些坐标 $x_j<y_j$。比方说，对某个 k 严格地成立不等式 $x_k>y_k$，即 $x_k>f_k(x)$，由于 $y_k=f_k(x)\geqslant 0$，此时必有 $x_k>0$，既然 $x_k\neq 0$，那么 $x_k>f_k(x)$ 一定不会发生在与 v^k 完全相对的边界上。

如果 $N=3$，我们已经知道单纯形是一个四面体，它有 4 个顶点、4 个边界面及 6 条边界线。6 条边界线段构成了 3 维单纯形的 1 维边界，而 4 个边界面实际上是 4 个三角形，它们构成了单纯形的 2 维边界。

很容易想象，N 维单纯形存在着 $(N-1),(N-2),\cdots,2,1$ 维边界。凡维数 $d<N$ 的边界显然由 $(d+1)$ 个顶点所确定。随便取一个 d 维边界，不妨设它的 $(d+1)$ 个顶点为

$$v^p,v^q,\cdots,v^s$$

由这 $(d+1)$ 个顶点界定的边界实际上是 d 维单纯形，以后由 v^p，$v^q,\cdots,v^s$ 为顶点的边界（或单纯形）均简记为 $\langle v^p,v^q,\cdots,v^s\rangle$。如果 $x\in\langle v^p,v^q,\cdots,v^s\rangle$，显然 x 的重心坐标必须满足

$$\begin{cases}x_p\geqslant 0,x_q\geqslant 0,\cdots,x_s\geqslant 0\\ x_k=0,k\neq p,q,\cdots,s\end{cases}$$

因此不等式 $x_k>f_k\geqslant 0(k\neq p,q,\cdots,s)$ 不可能发生在边界 $\langle v^p,v^q,\cdots,v^s\rangle$ 上。换句话说，假如在该边界上的点 x 对某些 j 成立 $x_j>y_j=f_j(x)$，那么 j 或者是 p，或者是 $q,\cdots$，或者是 s。

在完成上述预备工作后，现在我们对单纯形中的每一个点 x 定义该点的指标函数 $m(x)$：设 f 是 N 维单纯形到自身的映照。不妨设映照 f 不存在不动点（否则就无需证明不动点定理了），即 $x\neq y=f(x)$。这表明对于每一个点 x，总存在着某些下标 j 满足 $x_j>y_j$（当然也一定存在另外一些下标 k 满足 $x_k<y_k$，仅需考虑 $x_j>y_j$ 的情况即可）。这些下标 j 是 $0,1,2,\cdots,N$ 中的某几个，其中必存在一个最小的下标 j^* 使得 $x_{j^*}>y_{j^*}$。记 j^* 为

$$m(x)=\min\{j;x_j>y_j\} \qquad (3.10)$$

称 $m(x)$ 为点 x 的指标函数，其取值范围显然是 $0,1,2,\cdots,N$。正如前面所讨论的那样，在单纯形的边界上，$m(x)$ 必定受到约束，即，若 $x\in\langle v^p,v^q,\cdots,v^s\rangle$，则 $m(x)$ 只可能在 $p,q,\cdots,s$ 中取值。以 $N=2$ 为例，我们可以从图 3.2 中得到直观印象。

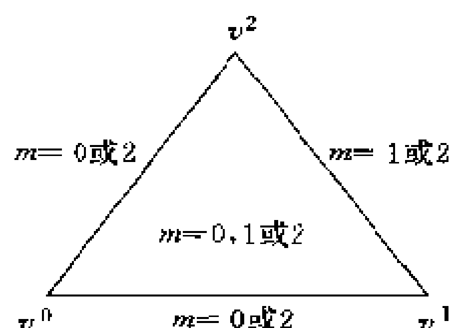


图3.2

由于我们假定 f 不存在不动点。事实上是试图借用反证法。为引出矛盾所采用的手段称为重心剖分。对于自然数 $m=2,3,\cdots$，我们可以对单纯形进行 m 次重心剖分，为更便于理解，不妨考虑 $N=2$、$m=2$ 这种最简单的情况。

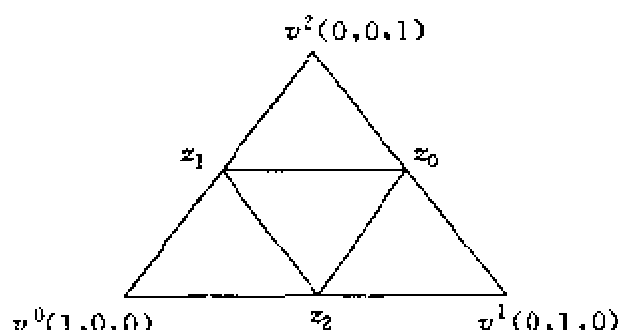


图3.3

在图 3.3 中，2 维单纯形（三角形）的三条边界分别被 z_0、z_1、z_2 等平均剖分，依据重心坐标原则，z_0、z_1、z_2 重心坐标分别为

$$z_0=(0,1/2,1/2)=1/2(0,1,1)$$

$$z_1=(1/2,0,1/2)=1/2(1,0,1)$$

$$z_2=(1/2,1/2,0)=1/2(1,1,0)$$

即二次剖分的顶点的重心坐标为

$$\left.\begin{array}{l} z=\frac{1}{2}(k_0,k_1,k_2) \\ k_i\geqslant 0 \quad (i=0,1,2)\text{为整数}, \quad \sum_{i=0}^{2}k_i=2 \end{array}\right\} \tag{3.11}$$

不难设想，如果 2 维单纯形的边界等分为 m 部分，然后将两条边的剖分点相连（必须使连线恰好平行于第三条边），这些连线将 2 维单纯形划分为 m^2 个小 2 维单纯形，共有 $(m+1)(m+2)/2$ 个剖分顶点（包括原来单纯形顶点在内），如图 3.4 所示。

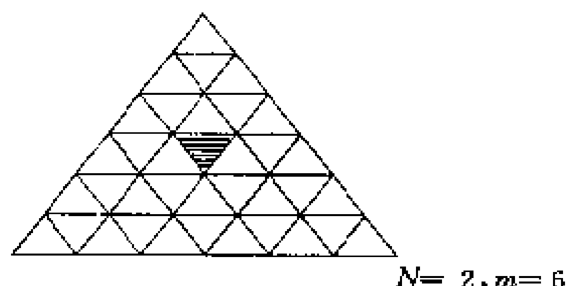


图3.4

这些顶点的重心坐标总可以写成（设单纯形为 N 维）

$$\left.\begin{array}{l} x=\frac{1}{m}(k_0,k_1,\cdots,k_n) \\ k_i\text{为非负整数，且}\sum_{k=0}^{N}k_i=m \end{array}\right\} \tag{3.12}$$

仍以 $N=2$ 为例，我们称图 3.4 中的阴影小单纯形为一个格子。显然，当 m 越来越大时，格子数 m^2 将趋于无穷。由于原单纯形是有限的，因此每个格子的直径（例如 $N=2$ 时可定义作外接圆的直径）、面积及周长等均随之趋于零。

由于假定 $y=f(x)\neq x$，因此单纯形中每一个 x 均有 $m(x)$ 与之匹配，于是我们可以用指标函数 $m(x)$ 对 x 进行归类并在图中进行标号。显然，在单纯形的内点 x 将标以 $0,1,\cdots,N$ 中的某一个数，但在边界上的点 x，其标号当然地受到约束，即在边界 $\langle v^p,v^q,\cdots,v^s\rangle$ 上，$m(x)$ 必须满足

$$m(x)=p\text{ 或 }q\text{ 或}\cdots\text{或 }s \tag{3.13}$$

现观察 2 维单纯形的 m 次重心剖分，共有 $(m+1)(m+2)/2$ 个顶点，那么多个顶点，它们的标号只能为 0 或 1 或 2，而在三条边界上的顶点标号还要受到(3.13)式的约束。譬如在三角形的底边——即由 v^0 与 v^1 为端点的线段上的剖分顶点只能取 0 或 2；在右边边界上的剖分顶点其标号只可能取 1 或 2。唯有作为三角形内点的剖分顶点其标号才有可能在 0、1、2 中取一。图 3.5 是 $N=2, m=4$ 的标号示意图。

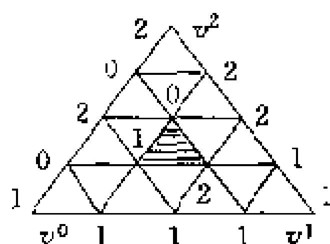


图3.5

在图 3.5 中，$N=2, m=4$ 重心剖分的顶点依据指标函数的定义及边界标号约束(3.13)式大致地被标上了号(具体标号其实应该与映照 f 有关，这里的标号纯属虚拟)。图中画有阴影的小格子的三个顶点，我们故意各标上 0、1 与 2 不同的数字(是否存在这种情况尚未知)，我们称这个小格子的顶点标号($\{0,1,2\}$)是完备集。现在要问，我们在图中虚拟标的完备集的小格子是否真的存在？一般地，如果是 N 维单纯形 m 次重心剖分的话，是否存在小格子其顶点的标号全体是 $\{0,1,\cdots,N\}$ 的完备集？为什么对这个问题会表示出如此特殊的兴趣，它对单纯形上不动点定理的论证提供了何种帮助，其理由详细阐述如下：

设 $y=f(x)$ 连续地将 N 维单纯形映射到自身，且总是 $x\neq y$ (即不存在不动点)。于是我们可以对每个点 x 定义指标函数，对 N 维单纯形的 m 次重心剖分所得各顶点可以根据指标函数及边界标号约束逐一进行标号。倘若对每一个 m，总是存在(至少一个)小格子，其顶点标号为完备集 $\{0,1,\cdots,N\}$ 的话，不妨将此小格

子的顶点及标号记作：

$$\left.\begin{array}{ll}\text{顶点 } x^0(m) & \text{标号为 } 0\\ \text{顶点 } x^1(m) & \text{标号为 } 1\\ \cdots\cdots & \\ \text{顶点 } x^N(m) & \text{标号为 } N\end{array}\right\} \tag{3.14}$$

根据指数函数与标号的定义，(3.14)式告诉我们，在顶点 $x^0(m)$ 必有 $x_0>y_0$，一般地，应有

$$\text{在顶点 } x^k(m)\text{，必有 } x_k>y_k(k=0,1,\cdots,N) \tag{3.15}$$

如 $N=2$ 情况一样，当剖分次数 $m\to\infty$ 时，小格子的直径、体积均趋于零，即小格子的所有 $(N+1)$ 个顶点越来越接近。因为

$$\max_{0\leqslant p<q\leqslant N}|x^p(m)-x^q(m)|=\Delta/m\to 0 \quad \text{当 } m\to\infty \text{ 时} \tag{3.16}$$

其中：Δ 表示小格子的直径。当 m 趋于无穷时，考虑序列

$$x^0(m) \quad m=2,3,\cdots$$

毫无疑问这一系列 $x^0(m)$ 均位于(有界的) N 维单纯形内，因为序列 $\{x^0(m),\ m=2,\ 3,\ \cdots\}$ 是有界序列。由数学分析中最基本的“有界序列必有收敛子序列”定理，$x^0(m)$ 具有收敛子序列，不妨记作

$$x^0(m_t)\to x^* \qquad \text{当 } t\to\infty \text{ 时} \tag{3.17}$$

其中 x^* 为所收敛的极限，它显然属于 N 维单纯形。当 $t\to\infty$ 时，由于(3.16)式，顶点将越来越接近，因此

$$x^p(m_t)\to x^* \quad (p=0,1,\cdots,N) \quad \text{当 } t\to\infty \tag{3.18}$$

注意到映照 $f(x)$ 的连续性，应有

$$f(x^p(m_t))\to f(x^*)=y^* \quad (p=0,1,2,\cdots,N) \text{ 当 } t\to\infty \tag{3.19}$$

现在(3.15)式中令 m 为 m_t，(3.15)式当然仍成立

$$\text{在 } x^k(m_t) \quad \text{有 } x_k>y_k \quad (k=0,1,\cdots,N) \tag{3.20}$$

或者说

$$\text{在 } x^k(m_t) \quad \text{有 } x_k>f_k(x^k(m_t)) \quad (k=0,1,\cdots,N) \tag{3.21}$$

由于每个点 x 的重心坐标当然连续地依赖于 x 本身，因此在(3.20)式或(3.21)式中，令 $t\to\infty$，可得

$$\text{在极限点 } x^* \text{ 处}, x_k^* \geqslant f_k(x^*) = y_k^* \quad (k=0,1,\cdots,N) \tag{3.22}$$

(3.22)式蕴含着

$$x^* \equiv y^* = f(x^*)$$

极限点 x^* 即为 $f(x)$ 的不动点。于是我们轻而易举地证明了不动点定理，关键在于需要证明对 N 维单纯形的任意 m 次重心剖分必定存在顶点标号为完备集的格子。这就是著名的 Sperner 引理。

§3.3 Sperner 引理

Sperner 引理：对 N 维非退化单纯形进行任意 m 次重心剖分，对产生的各个小顶点标上指标函数 $m(x)$，但在边界上的顶点标号受到(3.13)式约束。只要 m 为大于 2 的任意确定的自然数，总存在某个小格子其各顶点的标号恰为 $0,1,2,\cdots,N$ 的完备集。

我们试图用数学归纳法证明 Sperner 引理，而且归纳的结论比 Sperner 引理更向前一步，那就是将"总存在某个小格子具有完备标号集"改为"具有完备标号集的小格子数必为奇数"。为此，先引进一些基本概念与记号：

在 N 维单纯形中，如果小格子的顶点标号各为 $m_0,m_1,\cdots,m_N$，则称该格子属于 $(m_0,m_1,\cdots,m_N)$ 类型，标号中的数 m_i 的排列次序不影响到类型。$(m_0,m_1,\cdots,m_N)$ 有点类似于从 $(0,1,\cdots,N)$ 中作 N 次有放回抽样的结果，在 $(m_0,m_1,\cdots,m_N)$ 中各拥有多少个 0 多少个 1 乃至多少个 N 决定了类型的本质，每一个小格子具有维数为 $0,1,\cdots,(N-1)$ 的边界，0 维边界就是小格子的顶点，$(N-1)$ 维边界则称之为面。不管怎样，一个 k 维边界恰有 $(k+1)$ 个顶点，每个顶点的标号又构成了 $(k+1)$ 维向量 $(m_0,m_1,\cdots,m_k)$，因此

各类维数的边界也存在类型$(m_0,m_1,\cdots,m_k)$。给定了 N 维单纯形中的 m 次重心剖分，在 Sperner 引理所规定的标号规则下，我们感兴趣的是形如$(a,b,\ \cdots,c)$这种类型的边界共有多少个，这个数字记作 $F(a,b,\ \cdots,c)$。例如图 3.5 中，$F(0)=4$，$F(1)=6$，$F(2)=5$ 分别表示标号为 0、为 1、为 2 的顶点个数；$F(0,0)=2$ 表示两端点均标 0 的线段型边界的个数，同理可得 $F(0,1)=3$，等等；$F(0,0,1)$则表示三个顶点分别标为 0、0、1 的小格子的个数，在图 3.5 中，$F(0,0,1)=1$。

利用引进的计数函数 $F(a,b,\ \cdots,c)$，Sperner 引理断言，在 N 维单纯形的 m 次重心剖分中，$F(0,1,\cdots,N)$必为奇数。

利用数学归纳法：

先考虑 $N=1$，即单纯形是一段线段。

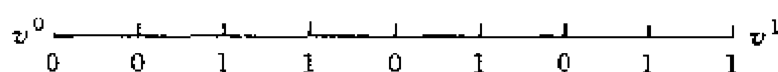


图3.6

如图 3.6 所示，$(N-1)$维边界就是图中剖分点。它们的标号只可能是 0 或 1。现在观察标号为 0 的剖分点，它们一定是(0,0)或(0,1)类型小格子的顶点。一个(0,0)型小格子拥有两个标号为 0 的顶点，而一个(0,1)型小格子只有一个标号为 0 的顶点。如果我们计算标号为 0 的剖分点个数为

$$2F(0,0)-F(0,1)$$

那么这个公式实际上将位于线段内部且又为(0,0)型小格子与(0,1)型小格子的共同边界点重复计算了一次，而且将两个相邻的(0,0)型小格子的共同边界也多计算了一次。因此，$2F(0,0)+F(0,1)$ 相当于线段内部标号为 0 的剖分点的个数的两倍再加上 v^0 这一点：

$$2F(0,0)+F(0,1)=2F_i(0)+1 \qquad (3.23)$$

其中，$F_i(0)$就表示标号为 0 的“内”顶点个数。(3.23)式蕴含了

$F(0,1)$必为奇数。即 Sperner 引理当 $N=1$ 时为真。

进而考虑 $N=2$。我们关心$(0,1)$类型边界的个数，这种类型的边界一定出现在如图 3.7 所示的几种类型小格子中。

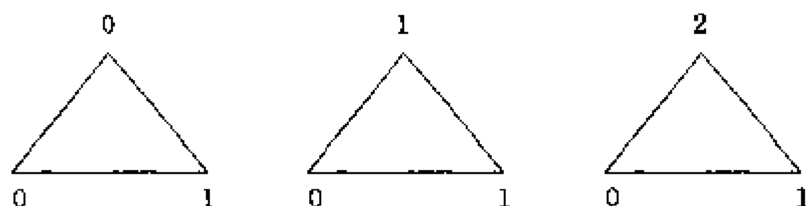


图3.7

图 3.7 中，左边与中间类型的小格子各拥有两条边界其类型是$(0,1)$型，而右边类型的小格子仅含有$(0,1)$型边界一条。如果以$2\{F(0,0,1)+F(0,1,1)\}+F(0,1,2)$来计算 $F(0,1)$的话，也存在位于单纯形内部的$(0,1)$型小格子边界被重复计算而位于 v^0 与 v^1 之间连线上的$(0,1)$型小格子边界肯定不会被计算两次的事实，若以 $F_i(0,1)$表示前者个数，$F_b(0,1)$表示后者个数，那么我们有

$$2\{F(0,0,1)+F(0,1,1)\}+F(0,1,2)=2F_i(0,1)+F_b(0,1) \tag{3.24}$$

由于 $F_b(0,1)$实际上是 $N=1$ 的情况，前面已证明 $F_b(0,1)$为奇数，因此由(3.24)式推出，$F(0,1,2)$为奇数。

现在不妨假设 Sperner 引理当 $N=k-1$ 时为真，即 $F(0,1,\cdots,k-1)$为奇数。继而考虑 $N=k$，对$(0,1,\cdots,k-1)$类型的边界面进行记数，这种类型一定出现在如下类型的小格子中：

$$类型(0,1,\cdots,k-1,t)\quad t=0,1,2,\cdots,k-1,k$$

显然$(0,1,2,\cdots,k-1,k)$型小格子仅含$(0,1,2,\cdots,k-1)$型边界面一个，而类型$(0,1,2,\cdots,k-1,t)$ $(t=0,1,2,\cdots,k-1)$的小格子均各拥有两个$(0,1,2,\cdots,k-1)$型边界面，因此总和

$$2\sum_{t=0}^{k-1}F(0,1,\cdots,k-1,t)+F(0,1,\cdots,k-1,k)$$

对位于 k 维单纯形内部的$(0,1,2,\cdots,k-1)$型边界面重复计算了

两次（因为它们是两个小格子的共同边界面），对以 $v^0,v^1,\cdots,v^{k-1}$ 顶点构成的边界面上 $(0,1,2,\cdots,k-1)$ 型小格子边界面仅计算一次（因为它们不可能是两个小格子的公共边界），于是我们有

$$2\sum_{t=0}^{k-1}F(0,1,\cdots,k-1,t)+F(0,1,\cdots,k-1,k)$$
$$=2F_i(0,1,\cdots,k-1)+F_b(0,1,\cdots,k-1) \qquad (3.25)$$

由归纳法假设，$F_b(0,1,2,\cdots,k-1)$ 必为奇数，因而 $F(0,1,2,\cdots,k-1,k)$ 亦必为奇数。我们已经证明了 Sperner 引理。

§3.4 Kakutani 不动点定理

通过 Sperner 引理实际上证明了 Brouwer 不动点定理，似乎在单纯形上可以推断 Nash 均衡的存在性。其实事情并非那么简单，因为在前面提及的反应函数（或对应）通常针对对手的每一个策略，局中人只有一个最佳反应，以致使我们所讨论的映照函数通常也被理解为单纯形上的一点通过映射可以在该单纯形中找到一点恰为它的映照。而我们试图证明的 Nash 均衡存在定理，策略空间包含了所有的混合策略，对于对手采取的混合策略 q（当然也可以是退化的混合策略，即纯策略），局中人需要选择混合策略 p，使自己的盈利（或效用）极大化。这样的 p 也许属于由若干概率向量所组成的某集合（即符合条件的 p 可以有若干个概率向量）。例如，局中人在知道对手的混合策略之后，计算自己采取混合策略 $p=(p_1,\ p_2,p_3,p_4)$ 所得的期望盈利

$$u=3p_1+5p_2+2p_3+5p_4$$
$$0\leqslant p_i\leqslant 1\quad(i=1,2,3,4),\ p_1+p_2+p_3+p_4=1 \qquad (3.26)$$

为使(3.26)式达到极大化，仅需注意到 p_2 与 p_4 前的系数相同且为最大，由于 $0\leqslant p_i\leqslant 1(i=1,2,3,4)$，且 $p_1+p_2+p_3+p_4=1$，因此 u 的值必小于等于 5，且在 $p_1=p_3=0$ 时达到 5，而无论 p_2 与 p_4 的

取值如何（只要 $p_2+p_4=1$ 即可）。也就是说，局中人使自己的期望盈利达到极大时的混合策略应当满足

$$\{p: p_1=p_3=0, 0\leqslant p_2, p_4\leqslant 1, p_2+p_4=1\} \tag{3.27}$$

显然，(3.27)式是若干个向量的集合，而不仅仅是一个概率向量。因此，我们所需要研究的映照应当是集值函数，而不仅仅是单值函数。

可见，为证明 Nash 均衡的存在性，有必要将 Brouwer 不动点定理拓宽到集值函数的范围。所谓集值函数 $Y=F(x)$，是指当 $x\in \mathscr{X}$时，Y 为$\mathscr{X}$中的一个子集。关于集值函数的不动点定理属于 Kakutani，它讨论如下关系：

$$X\in F(x) \tag{3.28}$$

意即当 $x\in \mathscr{X}$时，x 一定是$\mathscr{X}$的子集 $F(x)$中的一个元素。显然，如果子集 $Y=F(x)$只有一个元素 $y=f(x)$，那么(3.28)式又成为 $x=f(x)$，我们又回到了 Brouwer 不动点定理。

先叙述 Kakutani 不动点定理并作适当解释。

定理 3.2 (Kakutani)设$\mathscr{X}$是 N 维实空间 R^N 中的一个有界闭凸集，对于每一个 $x\in \mathscr{X}$，设 $F(x)$是$\mathscr{X}$中一个非空凸子集，假如"图"

$$\{x, y: y\in F(x)\} \tag{3.29}$$

是闭的，则存在 $x^*\in \mathscr{X}$，使得 $x^*\in F(x^*)$。

Kakutani 不动点定理是一条完全的纯数学定理，这里我们不去研究定理中所述条件是否必要或者是否有其他一些严密的陈述，我们要指出的是，若讨论的是完全信息静态有限博弈，其中的混合策略乘积空间 $\sum = X \sum_i$ 完全适合定理中的"$\mathscr{X}$是 N 维实空间 R^N 中的一个有界闭凸集"的要求。有限个局中人及有限纯策略空间保证了 $\sum$ 的有界性，而且如果将有限个纯策略表示为退化的混合策略（概率分布）的话，局中人的混合策略实际上是这些概率向量的凸组合，这个基本事实保证了所研究的策略空间的闭凸性。

重要的是解释(3.29)式关于“图$\{x,y:y\in F(x)\}$是闭的”这个许多人不太熟悉的概念：如果某序列 x^n 存在极限 x^0，又对于每一个 x^n，从集合 $F(x^n)$中取出一个元素记作 y^n，$y^n\in F(x^n)(n=1,2,\cdots)$。并假设这样获得的 y^n 也存在极限 y^0。“图$\{x,y:y\in F(x)\}$是闭的”则要求 $y^0\in F(x^0)$。事实上，这个定义讲述了集值函数 $F(x)$的上半连续性。

读者也许在直观上对这一概念缺乏理解。图的闭性，或者$F(x)$的上半连续性，实际上是通常意义下函数连续性的推广。设想每个子集 $Y=F(x)$只包含一个单点 $y=f(x)$，且在一有界闭集$\mathscr{X}$中，如果 $f(x)$是连续的，则图$\{x,y:y=f(x)\}$必定是闭的，因为$f(x)$的连续性保证了当 $x^n\to x^0$ 时必有 $f(x^n)\to f(x^0)=y^0$ 的缘故。反过来，如果 $f(x)$是上半连续的，那么 $f(x)$是否一定是连续函数呢？令 $x^n\to x^0$，由于$\mathscr{X}$是闭集，因此 $x^0\in\mathscr{X}$，记 $y^n=f(x^n)$ $(n=1,2,\cdots)$，且记 $y^0=f(x^0)$。现在我们需要证明 $y^n\to y^0(n\to\infty)$，即可得 $f(x)$的连续性。由于 $y^n=f(x^n)\in\mathscr{X}$，且$\mathscr{X}$是有界闭集，故由有界序列必存在收敛子序列这一著名的 Weirstrass 定理，存在 $n_i\to\infty$，使 $y^{n_i}\to y^*$。由于图的闭性，应有 $y^*=f(x^0)$，因此 $y^*=y^0$。利用同样的道理，我们可以证明 y^n 的所有收敛子序列都有同一极限 y^0，因此 y^n 收敛于 y^0。

最后对子集 $Y=F(x)$讲一点必须从数学角度讲的话，在 Kakutani 不动点定理中子集 $Y=F(x)$是非空凸子集这一假设是重要的，如果这个条件不满足，不动点定理可能不成立。例如设$\mathscr{X}=[-1,1]$，现定义集值函数

$$\begin{cases} F(x)=\left\{\frac{1}{2}\right\} & (-1\leqslant x<0) \\ F(0)=\left\{\frac{1}{2},-\frac{1}{2}\right\} & \\ F(x)=\left\{-\frac{1}{2}\right\} & (0<x\leqslant 1) \end{cases}$$

显然 $F(x)\subset\mathscr{X}$，且图 $\{x,F(x)\}$ 是闭的。但是在$\mathscr{X}$中没有一点满足 $x\in F(x)$，这里子集 $F(0)$ 包含两个点，它不是凸子集。

现在开始证明 Kakutani 不动点定理，方法仍借助于重心剖分。首先假设$\mathscr{X}$是 R^N 中的非退化单纯形，因为我们所介绍的重心剖分是在 N 维单纯形上进行的。假如$\mathscr{X}$不是单纯形怎么办呢？由定理假设，$\mathscr{X}$至少是有界的，因此总可以用一个单纯形 S 完全地包含$\mathscr{X}$。

以 $N=2$ 作为图示（见图 3.8）：

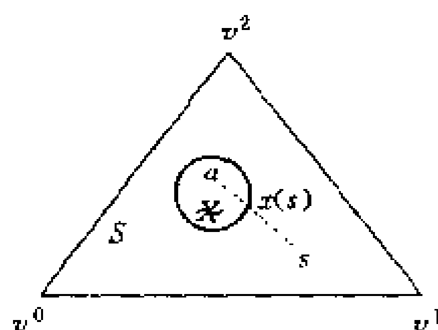


图3.8

可以将 S 连续地映射到$\mathscr{X}$，方法如下：

若 $s\in\mathscr{X}$，则 $x(s)=s$；

若 $s\in S$，但 $s\notin\mathscr{X}$，令 a 为$\mathscr{X}$一个内点，由于$\mathscr{X}$为有界闭凸集，因此内点 a 与 s 的连线上必有且仅有一个$\mathscr{X}$的边界点，此时则令 $x(s)$就是该边界点。

对每一 $s\in S$，定义子集

$$G(s)=F(x(s))\subset\mathscr{X}\subset S \tag{3.30}$$

由定理假设 F 是非空凸子集，故 G 也是凸的。现在证明图 $\{s,G(s)\}$ 是闭的，这是因为，假如

$$s\to s^0,\ y\in G(s)\quad 又\ y\to y^0 \tag{3.31}$$

那么

$$\left.\begin{array}{ll} x(s)\to x(s^0) & （由定义\ x(s)是连续映照）\\ y\in F(x(s))=G(s) & 以及\ y\to y^0 \end{array}\right\} \tag{3.32}$$

故有

$$y^0 \in F(x(s^0)) = G(s^0) \tag{3.33}$$

(3.33)式利用了定理中关于 $\{x, y: y \in F(x)\}$ 是闭的这一假设。

显然对于集值函数 $G(s)(s \in S)$，Kakutani 不动点定理的假设成立。如果 Kakutani 不动点定理关于单纯形为真的话。则存在 $s^* \in S$，使

$$s^* \in G(s^*) \tag{3.34}$$

由于 G 的定义，$G(s) = F(x(s)) \subset \mathscr{X}$，因此(3.34)式蕴含着 $s^* \in \mathscr{X}$。由 $x(s)$ 的定义可知，$x(s^*) = s^*$，因此，令

$$x^* = s^* = x(s^*) \in \mathscr{X} \tag{3.35}$$

显然有

$$x^* \in G(s^*) = F(x(s^*)) = F(x^*) \quad 即\ x^* \in F(x^*) \tag{3.36}$$

可见由于定理对单纯形成立导致它对凸集$\mathscr{X}$也成立。

因此我们仅需讨论$\mathscr{X}$为非退化的单纯形，不妨设

$$\mathscr{X} = \langle v^0, v^1, \cdots, v^N \rangle \tag{3.37}$$

构造$\mathscr{X}$的 n 次重心剖分。定义有关的连续映照 $f^{(n)}(x)$，令 x 为剖分小格子 $\langle x^0, x^1, \cdots, x^N \rangle$ 中的一点：

若 x 恰为顶点 $x^i(i=0,1,\cdots,N)$ 中的某一个，则从 $F(x^i)$ 中任取某点 $y^i \in F(x^i)$，并将它定义作 $y^i = f^{(N)}(x^i)(i=0,1,\cdots,N)$，又如果 $x \in \langle x^0, x^1, \cdots, x^N \rangle$，但不为其顶点，即

$$x = \sum_{i=0}^{N} \theta_i x^i \quad \theta_i \geqslant 0 \quad \sum_{i=0}^{N} \theta_i = 1$$

则定义

$$f^{(n)}(x) = \sum_{i=0}^{N} \theta_i f^{(n)}(x^i) \tag{3.38}$$

读者不必担心如果 x 恰好位于两个小格子的公共边界上这样的情况，例如位于 $(N-1)$ 维公共面上，(3.38)式是否会有不同的表示。因为(3.38)式告诉我们，公共边界上的点的 $f^{(n)}(x)$ 仅与边界的各顶点有关而与其他顶点无关。而 x 本身的表示也只与边界的顶点有关（有相同的顶点，对应地也就有相同的 θ_i），因此表示式是

一致的。

显然，如上述所定义的 $f^{(n)}(x)$ 连续地将单纯形 X 映射到自身。由 Brouwer 不动点定理，必存在一个不动点，记作 $x^{(n)}$，使得

$$x^{(n)}=f^{(n)}(x^{(n)}) \tag{3.39}$$

假如这个不动点恰好是剖分的某小格子的一个顶点，由 $f^{(n)}(x)$ 的定义，对于小格子的顶点，应有

$$x^{(n)}=f^{(n)}(x^{(n)})\triangleq y^{(n)}\in F(x^{(n)}) \tag{3.40}$$

(3.40)式表明此时 Kakutani 不动点定理成立。如果 $x^{(n)}$ 不是剖分的某小格子的顶点，现不妨设 $x^{(n)}$ 属于某格子 $\langle x^{n0},x^{n1},\cdots,x^{nN}\rangle$，设 $x^{(n)}$ 关于此格子的重心坐标为 $(\theta_{n0},\theta_{n1},\cdots,\theta_{nN})$，即

$$x^{(n)}=\theta_{n0}x^{n0}+\theta_{n1}x^{n1}+\cdots+\theta_{nN}x^{nN} \tag{3.41}$$

由于已证明 $x^{(n)}$ 是 $f^{(n)}(x)$ 的不动点，故

$$\begin{aligned}x^{(n)}&=f^{(n)}(x^{(n)})\\&=\theta_{n0}f^{(n)}(x^{n0})+\theta_{n1}f^{(n)}(x^{n1})+\cdots+\theta_{nN}f^{(n)}(x^{nN})\\&\triangleq\theta_{n0}y^{n0}+\theta_{n1}y^{n1}+\cdots+\theta_{nN}y^{nN}\end{aligned} \tag{3.42}$$

其中

$$y^{nj}=f^{(n)}(x^{nj})\in F(x^{nj})\ (j=0,1,2,\cdots,N) \tag{3.43}$$

(3.43)式来自于 $f^{(n)}(x)$ 在小格子顶点时的定义要求。再一次利用 Weierstrass 定理，即存在一个收敛的子序列，为记号上的方便起见，不妨令 $\{x^{(n)}\}$ 就是该收敛序列，于是当 $n\to\infty$ 时，应有

$$\left.\begin{aligned}&x^{(n)}\to x^{*}\\&\theta_{nj}\to\theta_{j}\qquad(j=0,1,\cdots,N)\\&y^{nj}\to y^{j}\qquad(j=0,1,\cdots,N)\end{aligned}\right\} \tag{3.44}$$

(3.44)式蕴含了 $x^{(n)}$ 所属的小格子 $\langle x^{n0},x^{n1},\cdots,x^{nN}\rangle$ 随 $n\to\infty$ 而趋于一点，即当 $n\to\infty$ 时

$$x^{nj}\to x^{*}\quad(j=0,1,\cdots,N) \tag{3.45}$$

且由于(3.44)式，我们得到的不动点 x^{*} 应有

$$x^* = \theta_0 y^0 + \theta_1 y^1 + \cdots + \theta_N y^N \tag{3.46}$$

由“图是闭的”这一假设，

$$y^j \in F(x^*) \quad (j=0,1,\cdots,N) \tag{3.47}$$

根据定理假设，子集 F 是凸的，所有 y^j 的凸组合必属于凸集 $F(x^*)$，因此

$$x^* \in F(x^*) \tag{3.48}$$

§3.5 Nash 均衡存在性

我们将利用 Kakutani 不动点定理来证明 Nash 均衡存在性。先叙述定理：

定理 3.3 (Nash 1950)任何有限正则型(或策略型)博弈具有混合策略均衡。

注：定理中有限策略型博弈的要求，指的是局中人的人数有限，且每一个局中人的纯策略空间含有有限个纯策略。其实在上一节对 Kakutani 不动点定理作解释时，我们已经指出，如果考虑的是有限策略型博弈，策略空间一定满足 Kakutani 不动点定理关于 $\mathcal{X}$ 的要求。

定理的证明：

不失一般性，令局中人人数 $N=3$，设

局中人 A 有纯策略 $i=1,2,\cdots,I$

局中人 B 有纯策略 $j=1,2,\cdots,J$

局中人 C 有纯策略 $k=1,2,\cdots,K$

对于每组纯策略组合(i,j,k)，记 A、B、C 三人的盈利为 a_{ijk}，b_{ijk}和 c_{ijk}，考虑混合策略：

A 取 i 的概率为 p_i，$(i=1,2,\cdots,I)$，$p=(p_1,p_2,\cdots,p_I)$

B 取 j 的概率为 q_j，$(j=1,2,\cdots,J)$，$q=(q_1,q_2,\cdots,q_J)$

C 取 k 的概率为 r_k，$(k=1,2,\cdots,K)$，$r=(r_1,r_2,\cdots,r_K)$

且假定 i、j、k 是独立的随机变量，局中人 A、B、C 的期望盈利分别为

$$\left.\begin{aligned} a(p,q,r) &= \sum_{i,j,k} a_{ijk} p_i q_j r_k \\ b(p,q,r) &= \sum_{i,j,k} b_{ijk} p_i q_j r_k \\ c(p,q,r) &= \sum_{i,j,k} c_{ijk} p_i q_j r_k \end{aligned}\right\} \tag{3.49}$$

若 A 知道对手所采取的混合策略 q 与 r，那么他一定适当地选取自己的 p 以使 $a(p,q,r)$ 极大化。从而 p 是 q、r 的函数，由前面的例子，这样的 p 可能不止一个，故可记作

$$p\in P(q,r) \tag{3.50}$$

如果引用记号

$$a_i=\sum_{j,k} a_{ijk} q_j r_k$$

那么

$$a(p',q,r)=\sum_i a_i p'_i \qquad p'\in P(q,r) \tag{3.51}$$

为使(3.51)式达到极大。p' 务必属于如下集合：

$$P=\{p:p_i\geqslant 0,\sum_i p_i=1,\text{如果 } a_i<\max_t a_t,\text{则 } p_i=0\} \tag{3.52}$$

(3.52)式清晰地表示集合 P 是一个有界闭凸集。由于 $a_1,a_2,\cdots$ 是 q、r 的函数，因此凸体 P 是 q、r 的集值函数。同样的讨论适用于对局中人 B 与 C。于是可以用

$$P(q,r),Q(p,r),R(p,q)$$

分别构成 A、B、C 三个局中人的最优混合策略集，其中 P、Q、R 均为有界闭凸集。如果

$$p\in P(q,r),q\in Q(p,r),r\in R(p,q) \tag{3.53}$$

则(p,q,r)是博弈的 Nash 均衡解。我们将证明有限策略型博弈中存在这样的解。

考虑 $I+J+K$ 维向量

$$x=\begin{pmatrix}p\\q\\r\end{pmatrix} \tag{3.54}$$

由于 $\sum_i p_i=\sum_j q_j=\sum_k r_k=1$，故向量 x 全体构成有界闭凸集$\mathscr{X}$。对于每一个 $x\in\mathscr{X}$，定义$\mathscr{X}$的子集

$$F(x)=\begin{pmatrix}P(q,r)\\Q(p,r)\\R(p,q)\end{pmatrix} \tag{3.55}$$

P、Q 与 R 已证为凸体。Nash 定理无非是要证明至少有一点 x^* 属于它自己 $F(x^*)$，或者说，$p^*\in P(q^*,r^*)$，$q^*\in Q(p^*,r^*)$ 及 $r^*\in R(p^*,q^*)$。这与 Kakutani 不动点定理的结论是一样的，因此关键在于要验证 Kakutani 定理的假设是否满足。即验证

(1)子集 F 在$\mathscr{X}$中是凸的；

(2)图$\{x,F(x)\}$是闭的。

条件(1)是由 $F(x)$的定义自然满足的，因为 P、Q、R 为凸集，而显然 $F(x)\subset\mathscr{X}$。

现在来验证条件(2)，令

$$x_n\in\mathscr{X}\quad y_n\in F(x_n) \tag{3.56}$$

并设这些 x_n 与 y_n 分别在闭集$\mathscr{X}$中有级限

$$x_n\to x_0\in\mathscr{X}\qquad y_n\to y_0\in\mathscr{X} \tag{3.57}$$

关键是需要证明

$$y_0\,\mathscr{X}\in F(x_0) \tag{3.58}$$

为此，令

$$x_n=\begin{pmatrix}p_n\\q_n\\r_n\end{pmatrix}\quad y_n=\begin{pmatrix}u_n\\v_n\\w_n\end{pmatrix}\tag{3.59}$$

由 y_n 的定义，它来自 $F(x_n)$中的某一点，因此再由 $F(x)$的定义易知 $u_n\in P(q_n,r_n)$，$v_n\in Q(p_n,q_n)$，$w_n\in R(p_n,q_n)$。再由 P、Q、R 的定义，对任意向量 p' 有

$$a(u_n,q_n,r_n)\geqslant a(p',q_n,r_n)\tag{3.60}$$

对任意向量 q'，有

$$b(p_n,u_n,r_n)\geqslant b(p_n,q',r_n)\tag{3.61}$$

对任意向量 r'，有

$$c(p_n,q_n,w_n)\geqslant a(p_n,q_n,r')\tag{3.62}$$

当 $x_n\to x_0$，$y_n\to y_0$（当 $n\to\infty$），其分量 p_n、q_n、r_n与 u_n、v_n、w_n 应各自趋于自己的极限 p、q、r 与 u、v、w。注意到盈利函数的连续性（见(3.49)式关于 a、b、c 的定义），在(3.60)式、(3.61)式与(3.62)式的两端令 $n\to\infty$，取极限，易得

$$\left.\begin{aligned}a(u,q,r)\geqslant a(p',q,r)\\ b(p,u,r)\geqslant b(p,q',r)\\ c(p,q,w)\geqslant c(p,q,r')\end{aligned}\right\}\tag{3.63}$$

由于 p'、q'、r'的任意性，可见

$$y_0=\begin{pmatrix}u\\v\\w\end{pmatrix}\in F(x_0)$$

既然 Kakutani 定理的假设条件都成立，那么三人有限博弈必存在解 x^*：$x^*\in F(x^*)$。

当 $N>3$ 时，证明过程完全一样，只不过略繁复一些。

我们来总结一下有限策略型博弈存在 Nash 均衡的证明。我们是先从单纯形上的 Brouwer 不动点定理开始的（其实 Brouwer

不动点定理是对有界闭凸集合而言的，由于拓扑等价的缘故，我们仅需在非退化单纯形上进行证明即可，一般非数学专业的读者不必在意，仅需知道这一事实），然后我们利用 Brouwer 不动点定理证明 Kakutani 不动点定理，它们之间的差异主要在于后者在有界闭凸集上将映射拓宽到集值函数，这一拓宽更与博弈模型雷同。借助于 Kakutani 不动点定理，我们证明了有限个局中人的有限（正则型）博弈中的 Nash 均衡存在性，在证明过程核实条件时，关键的一步是证实当盈利函数是连续函数时，反应对应有闭图。注意到盈利函数的连续性（见(3.49)式）与我们所考虑的空间是混合策略空间有关。而混合策略空间又保证了 $\sum$（即$\mathscr{X}$）的有界闭凸性等 Kakutani 定理的条件。这表明 Nash 均衡存在性在有限正则型博弈中只有在混合策略空间才成立，在纯策略空间由于无法保证$\mathscr{X}$的有界闭凸性也不能保证盈利函数的连续性等条件，因此博弈有可能不存在纯策略 Nash 均衡。猜谜博弈是一个很能说明问题的例子，它没有纯策略解，但是却有混合策略解。

§3.6 连续盈利无限博弈中的 Nash 均衡存在性

我们已经研究了有限策略型博弈的 Nash 均衡存在性问题。当行动空间或策略空间呈无限不可数状态时，我们已经接触到 Cournot 竞争博弈，在那里我们假设行动空间具有连续统。经济学家或许会争辩，在社会经济与管理领域，大量博弈模型存在着无限多个行动，但策略或行动空间并非一定具有连续统。例如 Cournot 模型中策略为产量，Bertrand 模型中局中人选择的价格。价格或某些产量实际上是离散的，连续性只不过是数学上的提炼，当然正如我们在第二章中所提到的，由“厘”、“毫米”这些数学上的微小单位构成的格子，事实上如此地稠密，以至我们如果近似地将策略空间视作连续的，在处理时将增加不少的方便。可是有一个理论上的问

题摆在我们的面前，假如诸如价格那样的行动空间范围有限，那么小格子即使再密，行动空间仍为有限的，有限博弈必定存在着 Nash 均衡，而一旦将策略空间处理为具有连续统，Nash 均衡的存在性是否会成问题。事实上可以设想到，倘若将策略空间视作具连续统去处理时不存在 Nash 均衡，那么策略空间里那些稠密且离散的格子(此时却存在 Nash 均衡)的不同划分使所对应的 Nash 均衡必定相当敏感：同一博弈中格子的不同选择可能对应截然不同的 Nash 均衡。这个事实由 Dasgupta 与 Maskin 在 1986 年讨论过。因此我们可以继续设想、在经济博弈问题中，假如存在有限格子型博弈的均衡，且对于格子的选择相当地不敏感，如通常微积分中处理问题那样，我们可以取一系列"收敛"于连续统的越来越精细的格子，相应的离散行动空间的均衡的收敛子序列的极限无疑将是近似连续假设下的均衡。

本节主要考虑的是策略空间真正具有连续统时正则型博弈的 Nash 均衡存在性。有如下定理：

定理 3.4 (Debrreu 1952, Glickssherg 1952, Fan 1952)考虑一个策略型博弈，其中各局中人的策略空间 S_i 为欧氏空间中非空紧凸子集，倘若盈利函数 U_i 关于策略剖面 s 为连续且关于局中人 i 的纯策略 s_i 为拟凹，那么博弈存在一个纯策略 Nash 均衡。

我们将对定理中提到的若干概念作一些详细说明之后再证明定理本身。这些叙述是面对非数学专业的经济应用人员，因此不太讲究严格性。

考虑 n 维欧氏空间 R^n，R^n 的一个子集 C 称为凸的，如果 $x,y\in C, 0\leqslant\alpha\leqslant 1$，则有 $\alpha x+(1-\alpha)y\in C$。直观地描述就是指子集中任意两点的连线上所有点均属于该子集。R^n 上任意子集 S 并不都是凸的，如果将 S 上任意两点连线顶点全包括在一起则构成了

的凸包。

一个集合称为紧或紧致，是指集合中任意一列收敛子序列的极限均属于该集合。读者不难想象：R^n 中有界凸集如果是紧的，那么其边界一定属于该凸集，因此是闭的。

现在需要介绍定理中一个重要概念——拟凹函数：

令$\mathscr{X}$是 n 维欧氏空间 R^n 中的一个凸子集，且令 f 为从$\mathscr{X}$到实数空间 R 的函数。函数 f 称为在$\mathscr{X}$中为拟凹的(quasi-concave)，如果对每一个 $r\in R$，集合

$$\{x\in \mathscr{X} \mid f(x)\geqslant r\} \tag{3.64}$$

为凸的。或者等价地：如果对任意 x，$y\in \mathscr{X}$，以及$[0,1]$区间中任意的 t，成立

$$f(tx+(1-t)y)\geqslant \min[f(x),f(y)] \tag{3.65}$$

注意拟凹函数与凹函数之间的差异，如果(3.65)式成为

$$f(tx+(1-t)y)\geqslant tf(x)+(1-t)f(y) \tag{3.66}$$

则称函数 f 为凹函数，倘若(3.66)中不等号严格成立，则 f 被称为严凹函数。由于 $t\in[0,1]$，故

$$tf(x)+(1-t)f(y)\geqslant \min[f(x),f(y)] \tag{3.67}$$

也就是说凹函数必定是拟凹函数。

定理 3.4 假定 U_i 关于 s_i 为拟凹，实质上是假设了反应对应是凸映射(即集值函数 F 在$\mathscr{X}$中是凸的)。这是因为拟凹性保证了如果 s_{i1}与 s_{i2}是关于其他对手策略的最佳反应的话，由于对任意的 $\lambda\in[0,1]$，有

$$u_i(\lambda s_{i1}+(1-\lambda)s_{i2})\geqslant \min[u_i(s_{i1}),u_i(s_{i2})] \tag{3.68}$$

因此 $\lambda s_{i1}+(1-\lambda)s_{i2}$亦是最佳反应。

前面已经指出，盈利函数的连续性实际上保证了反应对应的闭图要求。一旦解决了这两个问题，在定理 3.4 证明中验证 Kakutani不动点定理的条件就成了相当简单的事情，仿照 Nash 均衡存在性定理的证明不难获得定理 3.4，只不过在 Nash 定理中

的策略空间为混合策略全体，从而保证了X为非空紧凸集的要求，而在这里的策略定理则为欧氏空间中的非空紧凸子集，因此按不动点定理得到的均衡点实际上是纯策略 Nash 均衡。

盈利函数的连续性是非常重要的条件。倘若它不是连续的，反应对应就可能没有闭图或者不是非空的。为什么可能不是非空的呢？那是因为不连续函数不一定达到最大值。最简单的例子是 $f(x)=-|x|$（当 $x\neq 0$ 时），$f(0)=-1$。

现在再举例说明盈利函数的不连续性引起反应对应没有闭图的可能性。考虑如下双人博弈：

$$S_1=S_2=[0,1]$$

$$u_1(s_1,s_2)=-(s_1-s_2)^2$$

$$u_2(s_1,s_2)=\begin{cases}-\left(s_1-s_2-\dfrac{1}{3}\right)^2 & s_1\geqslant\dfrac{1}{3}\\ -\left(s_1-s_2+\dfrac{1}{3}\right)^2 & s_1<\dfrac{1}{3}\end{cases}$$

$u_2(s_1,s_2)$在 $s_1=1/3$ 处间断。可是 u_1 关于 s_1，u_2 关于 s_2 显然是严格的凹函数，因此对于对手的每一策略。必定存在唯一的最佳反应，但是该博弈没有纯策略均衡：可以计算得局中人 1 的反应函数为 $r_1(s_2)=s_2$，而局中人 2 的反应函数为当 $s_1\geqslant 1/3$时，$r_2(s_1)=s_1-1/3$；当 $s_1<1/3$ 时，$r_2(s_1)=s_1+1/3$。这两条对应曲线不相交。

拟凹性在定理中是充分条件而不是必要条件，即使在有些场合不满足，不等于博弈不存在 Nash 均衡解（纯策略的）。例如 Cournot 博弈中，如果盈利函数不是我们例举的简单线性，那么拟凹性对价格或产量将有较高的要求，诸如二阶导数方面的要求等等。

盈利函数的拟凹性在定理证明中起到保证凸映射的重要作用，但它又不是必要条件，如果我们将拟凹性去除又怎么样呢？看来要保证存在纯策略 Nash 均衡可能会成问题，因为就纯策略而

言，此时已经无法保证反应为凸映射了。但倘若我们引入混合策略，仍然可以得到凸反应。Glicksberg 于 1952 年得到了下述结果；

定理 3.5 (Glicksberg 1952)考虑策略型博弈，其局中人的策略空间 S_i 是度量空间中的非空紧子集，如果盈利函数 u_i 为连续函数，那么博弈至少存在着一个混合策略的 Nasb 均衡。

这里我们不再重复叙述定理的证明过程，实质上是再一次地逐一验证 Kakutani 不动点定理的各种条件。需要强调一下，在具连续统集合的纯策略空间上的"混合策略"实际上是纯策略空间上的概率分布。例如局中人 1 的纯策略空间是 $S_1=[0,1]$，那么他的混合策略空间是[0,1]上的概率分布的全体。

第二部分　完全信息动态博弈

在完全信息静态博弈中，我们研究了局中人“同时”行动的可能结局，即使局中人采取行动的时间不一定一致，但每个局中人都不知道也无法观察到其他局中人的行动。假定在时间上稍晚一些行动的局中人可以观察到在他之前其他局中人采取的行动，出于理性，此时他可能的选择常常比静态博弈时的选择面小一些——他可以“有条件”地选择自己的策略。譬如在囚徒窘境中，若乙囚徒获知甲囚徒（肯定地）拒绝坦白，那么他很可能也拒绝坦白，以使博弈达到有效结局。假如“先行”的甲囚徒能预测到乙在看到自己拒绝的态度之后也会同样强硬地抗拒从而使双方都得到好处，他必将毫不犹豫地抗拒。显然，一旦静态转为动态，博弈的结局有可能发生根本性的改变。生活中的动态博弈比静态博弈多得多，最常见的例子是下棋。后走棋者通常根据先走棋者的下法考虑对策，而先行的棋手在自己走出一步棋之前往往要考虑到对手将作出何种反应。

在动态博弈中存在着许多静态博弈所看不见想不到的现象与问题有待我们研究。比如，由于动态博弈中的局中人的行动有先有后，后行动者又能观察到先行动者的行为，其间就会产生一个可信度（credibility）问题。后行动者可以“承诺”采取对先行者有利的行动，也可以“威胁”先行者以使先行者不得不采取对后行动者有利

的策略。“承诺”与“威胁”就存在一个可信不可信的问题。例如，局中人 2 要局中人 1 拿出 1 万元给他，否则将引爆炸弹同归于尽。这是一个威胁，如果局中人 1 相信威胁是真的，他会乖乖地拿出钱而结束博弈。如果局中人 1 认为局中人 2 本身很怕死，因此威胁是空的、不可信的，博弈的结局自然可能是另一种状态。当然，局中人 1 可以怀疑局中人 2 是疯子，在吃不准局中人 2 究竟是否疯子时，实际上在博弈模型中可以利用一个概率分布进行建模。吃不准对方是否为疯子的局中人 1 处于一个不完全的信息集中，承诺与威胁是否可信构成动态博弈的中心问题之一。

动态博弈与静态博弈的不同之处还在于动态博弈存在着子博弈，即从某一阶段以后局中人的一系列对策与行动直至博弈结束的整个博弈过程。形象的例子是下棋中的残局。一个棋手赢了一盘棋，如果他的整个下法在这盘棋中从每一步开始的残局里都称得上最优的，人们可能会称赞他的这局棋达到完美无缺的地步。在动态博弈中，也存在着类似的“理想”结局，这就是我们将要介绍的“子博弈完美均衡”——这是博弈论中一个极其重要的基本概念。

动态博弈存在着先后行动，在局中人每一次进行决策之前，他可以根据所观察到的以前行动，或者说根据自己所获知的信息作出自己的最佳选择。每个局中人都这样一步一步地展开对策或较量，直至博弈结束。因此它不像静态博弈那样仅仅“一次性地较量”，而是有环节或有阶段或有层次地将博弈过程展开下去。因此，动态博弈具有独特的展开型表述方式，我们将在第四章对所谓展开型博弈进行讨论。

第四章　展开型博弈

§4.1　定义与博弈树

博弈的展开形式包含下述信息与内容：

(1)局中人的集合，仍记为$i=1,2,\cdots,I$。这与静态情况完全一样，局中人当然地构成博弈的最基本要素。

(2)行动的次序，即谁在什么时候行动。

(3)当一个局中人行动时他的选择是什么；实际上就是轮到他行动时，他从该时刻的纯策略空间中选取了什么策略。

(4)当局中人作出他的行动决策时，他所观察到或他所了解的东西，这就是他在此时获得的信息集合。

(5)局中人的盈利或效用，它们是已采取行动的函数。像静态时一样，盈利函数构成博弈的基本要素。

(6)在任何外生事件上的概率分布。

例如，动态博弈在天是否下雨的条件下以不同方式展开，这里天是否下雨是博弈的外生事件，关于晴天或者下雨的概率分布可以认为是"自然"采取的行动，局中人"自然"常常用N来表示。

正因为行动具有次序，可以依次序将一步又一步的行动展开成图形。我们称该图形为博弈树(game tree)，它有效地向人们展示了局中人的行动、选择这些行动的次序以及作出决策时局中人

所拥有的信息集。博弈树是由结(nodes)与枝(branches)组成的图。博弈树中的每一个结表示某局中人(不妨设 i)的决策点，并称此点属于在该点行动的局中人。常在该点上方(或旁侧)标上行动局中人的代号 i。枝则表示局中人可能的行动，每一个枝连接两个结并且具有从一个结到另一个结的方向，常用箭头来表示。如果枝是从属于局中人 i 的结 N_1 到属于局中人 j 的结 N_2，那么动态博弈中局中人 i 行动在局中人 j 之前，并称结 N_1 直接位于结 N_2 之前。在所有的结中，有两种特殊的结：初始结与终点结。初始结(initial node)即在它之前没有任何其他的结，它表示由"自然"开始的行动(如果存在的话)，倘若"自然"的行动是平凡的，则表示由某个局中人(例如局中人 1)开始的行动。这是整个动态博弈的出发结，许多书中常用空心圆"○"来表示。终点结则是没有任何后续结(即位于它后面的结)的结，它实际上表示博弈的结束，因此在该结点没有任何局中人行动，在它的上方或旁侧就没有标上任何局中人代号，不少书中干脆连黑点"●"(即一般的决策结)也不绘出，但由于此时表示博弈结束，各局中人通过博弈到达终点结时各有所获，于是在终点结那儿会出现盈利(或效用)向量$(u_1, u_2, \cdots, u_I)$分别表示各局中人的获益。

为了避免不明确性，博弈树必须服从四条规则：

博弈树规则一：每一个结至多有一个其他结直接位于它的前面。

规则一排除了如图 4.1 中的情况。

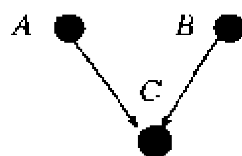


图4.1

我们希望博弈树中排除图 4.1 这样的情况是因为在考虑模型中，博弈树的每一个结意指在它之前发生的所有事件的全部描述。

当规则一被满足时，谈论一个决策结跟在另一个结的后面才有意义。非正式地，如果可能的话，称结 B 为结 A 的后继者，是指博弈从 A 出发，这些局中人作出一系列行动以使博弈达到结 B。正式地，结 B 称为结 A 的一个后继者，当且仅当存在某系列结，N_1，N_2，…，N_k 使得 $A=N_1$，$B=N_k$，每一个结位于在该系列中下一个结的前面。该系列结被称为从 A 到 B 的一条路径(path)。规则一意思是说在博弈树的两个结之间至多只有一条路径(注意：在叙述中我们使用了决策结(decision node)这个术语，其实非终点结的那些结都是决策结)。

博弈树规则二：在博弈树中没有一条路径可以使决策结与自身连接。

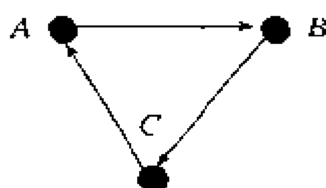


图4.2

规则二事实上避免了具有循环图的决策树，从图形上看，排除了如图 4.2 中的不明确情况。在图 4.2 中出现了一个很大的问题：局中人 A、B、C 三人到底谁先行动？显然该情况违反了博弈树规则二。

博弈树规则三：每个结是唯一的初始结的后续者。也就是说，博弈树必须有初始结。

我们研究的动态博弈是从最初时刻开始逐步展开的，因此总是要求博弈树具有初始结。

博弈树规则四：每个博弈树正好"只"有一个初始结。

规则四并不意味着所有的博弈树不会发生多于一个初始结的情况，但是在一般情况，倘若发生了两个及两个以上的初始结，我们常常可以将它们分解为若干博弈树，或者利用"自然"构成一个

限于这几个初始结的“原始初始结”，因此在展开型博弈中我们总是假设初始结是唯一的。

根据上述博弈树的规则，对于博弈树中的每一个终点结，我们可以完全确定从初始结到该终点结的一条路径，同时也展示了到达该结局的博弈动态过程。动态博弈中很重要的一条是当局中人选择自己的行动时所拥有的信息。我们利用信息集及记号 $h\in H$ 来表示这个信息。信息集把博弈树的结分成若干部分，每一个结恰好在一个信息集上。如果一个信息集包含结 x，我们可以将该信息集记作 $h(x)$，假使一个信息集只包含一个结，那么这是最简单且特殊的情况，稍后立即会涉及这种情况。我们关心的是一个信息集不止只包含一个结，假设 x 与 x' 都属于同一个信息集 $h(x)$，那么恰好拥有信息 $h(x)$ 并正在选择自己行动的局中人其实对自己究竟处于结 x 或处于结 x' 是不确定的。我们要求，如果 $x'\in h(x)$，则在 x 与 x' 应该由同一个局中人采取行动。没有这个要求，博弈的局中人在让谁行动这一点上可能发生争执。同时，在上述要求下我们进而要求，如果 $x'\in h(x)$，那么在结 x 与结 x' 上局中人可选择的策略空间是相同的，记为 $A(x)=A(x')$。因此可以将信息集 h 上的行动集干脆记作 $A(h)$。

假如一个局中人在轮到他行动时知道自己处于博弈树的那个结上，我们称局中人有完美信息(perfect information)。博弈中的每一个局中人都具有完美信息，则称该博弈有完美信息，下棋就是具有完美信息的一个例子。反过来，如果局中人在不知道另外的局中人在前面究竟怎样行动的情况下必须行动，则称该局中人具有不完美信息(imperfect information)。倘若至少有一个局中人具有不完美信息，则称博弈具有不完美信息。如果一个局中人 i 在行动时具有不完美信息集 $h(x)$，$h(x)$ 包含结 x、x'，…，y。那么在博弈树上体现为用一根虚线将这些结连接起来。虚线上有时标上该局中人的代号 i。假如我们将这个概念与完全信息静态博弈联系起

来，由于局中人同时行动，每个局中人都不知道对手采取什么行动，因此静态博弈具有不完美信息。

完全信息与完美信息是不能混为一谈的两个概念，完全信息是指盈利函数和纯策略空间均为博弈各方的共同知识，完全信息可以是完美的也可以是不完美的，轮到行动的局中人不知道先前行动（或同时行动）的其他局中人采取了什么策略，此时他拥有的信息集就是不完美的。

我们现在对产业组织理论中的"进入与阻扰"模型依据上述规则绘出博弈树。

例 4.1 设局中人 1 为某产品市场的"在位者"，局中人 2 打算是否进入该市场，如果局中人 2 作出"进入"的决策，那么双方之间形成竞争，例如在产量方面各有两种选择，我们在省略盈利函数情况绘出如图 4.3 所示的博弈树。

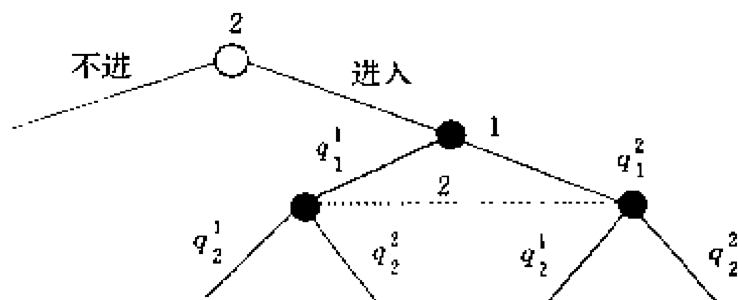


图4.3

注 1：图 4.3 的初始结属于局中人 2，没有人规定初始结必须从局中人 1 开始，但倘若存在"自然的行动"，那么无可非议地初始结属于自然。

注 2：作为博弈树，图 4.3 是自上而下展开的，有些书上也出现博弈树的展开为自左向右，这两种情况都是博弈论研究人员认可的习惯。

§4.2 展开型博弈的策略与均衡

1. 行为策略(behavior strategies)

正则型博弈中，局中人的策略是进行博弈的计划（或打算）的

详细集合，沿用策略的这个意思，展开型博弈中局中人的策略必须确定在该局中人的每一个决策结上所采取的行动。我们知道，结与信息集紧密连接，因此，对于局中人 i，基于信息集 h_i 的行动全体记为 $A(h_i)$，如果令 H_i 表示局中人 i 的信息集的集合的话，那么 $A_i=U_{h_i\in H_i}A(h_i)$就是局中人 i 的所有行动的集合。局中人 i 的一个纯策略是从 H_i 到 A_i 的一个映射 s_i：即对一切 $h_i\in H_i$，$s_i(h_i)\in A_i$。由于从同一个信息集出发，i 可以取 $A(h_i)$（或 A_i）中的不同行动，因此 $s_i(h_i)$有不同行动，也即有不同的纯策略 s_i，所有这些 s_i 的全体构成了局中人 i 的纯策略空间 S_i，根据 s_i 与 S_i 的定义，我们可以记 S_i 为在每一个信息集 h_i 的行动空间 $A(h_i)$的笛卡儿乘积空间(Cartesian Product)：

$$S_i=\underset{h_i\in H_i}{X}A(h_i) \tag{4.1}$$

(4.1)式对只关心博弈论的经济应用的读者也许在理解方面造成困难。我们用下述简单的展开型博弈作一些直观理解(见图 4.4)。

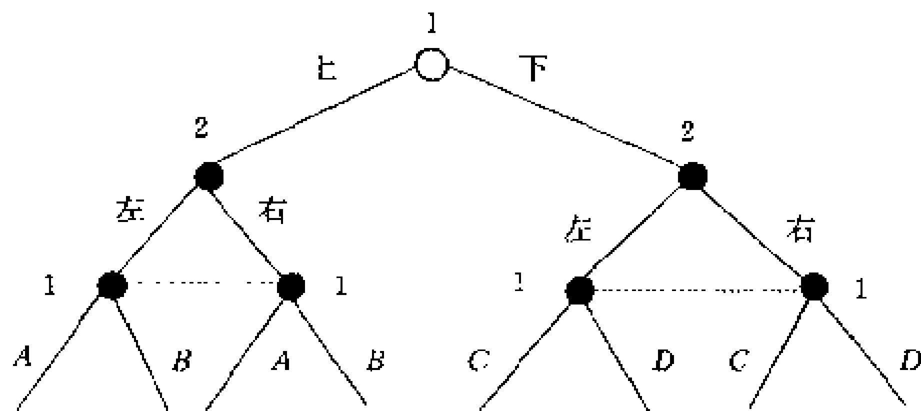


图4.4

图 4.4 中，局中人 2 有两个决策结，因此也有相应的两个信息集：局中人 1 取“上”或“下”。这两个信息集的行动空间均包含“左”、“右”两个行动。即 $A(h_2^1)=A(h_2^2)=\{左，右\}$，(4.1)称局中人 2 的纯策略空间为

$$S_2=(A(h_2^1),A(h_2^2))$$

$=\{(左,左)(左,右)(右,左)(右,右)\}$ (4.2)

纯策略(左,左)表明“当局中人 1 取‘上’时局中人 2 取‘左’,当局中人 1 取‘下’时局中人 2 取‘左’”;类似地,纯策略(左,右)表示“局中人 1 取‘上’时局中人 2 取‘左’,而当局中人 1 取‘下’时局中人 2 取‘右’”;同样可以推得纯策略(右,左)与(右,右)的确切含义。S_2 的个数相当于 $A(h_2^1)$的个数 2 乘以 $A(h_2^2)$的个数 2,总共有四个纯策略。按照纯策略空间的定义(4.1)式。我们可以知道局中人 1 的纯策略可以有(上,A,C)、(上,A,D)、(上,B,C)、(上,B,D)、(下,A,C)、(下,A,D)、(下,B,C)、(下,B,D)等 8 种。

一般地局中人 i 的纯策略空间的纯策略数目(这里我们暂时考虑有限情况)为

$$\# S_i=\prod_{h_i\in H_i}\#(A(h_i))\qquad(4.3)$$

记号 $\# C$ 表示集合 C 中所含元素个数。展开型博弈中纯策略是由信息集与行动集定义的,但策略与行动并不是等同的事情。这一点与静态博弈不同(在静态博弈中,采取某纯策略与采用某行动的意思是一样的)。不同的策略可以使局中人发生相同的行动,例如,设想图 4.4 中的博弈,当局中人 2 采取行动后就宣告博弈结束(即最后局中人 1 的决策结改为终点结),那么不同的策略剖面{上,(左,右)}与{下,(左,左)}均导致局中人 2 采取相同的行动“左”,但是最后结局是不同的。

博弈的纯策略剖面(pure strategies profile)是由局中人各自纯策略空间中的任一纯策略构成的剖面。由纯策略的定义,在任一纯策略剖面 s 下,总是可以从初始结开始,沿着博弈树的某条路径达到 s 相应的终点结。有一个事实非常重要:s 中有些信息集在博弈树的这条路径上,我们称这些信息集是 s 的路径(path),当然也可能存在 s 中的某些信息集不在此路径上。在许多经济应用中,初始结常从“自然”开始,“自然”以一定概率从初始结出发选择其行动,或者说,在“自然”的选择行动上具有一定的概率分布(在实际

操作的博弈树图中，常在从“自然”的初始结引出的枝旁用中括号标出相应概率）。根据这个概率分布以及每一个局中人 i 的纯策略，完全可以计算出在终点结（或结局）上相应的概率分布，并对每一个纯策略剖面 s“指定”局中人 i 的期望盈利 $u_i(s)$。事实上，s 路径上的信息集总是以正概率达到，而 s 中那些不在路径上的信息集达到的概率为零。

定义了纯策略的盈利函数之后，我们就可以定义展开型博弈的纯策略 Nash 均衡为满足如下要求的纯策略剖面 s^*：在给定局中人 i 的对手的策略 s_{-i}^* 时，每一个局中人 i 的策略 s_i^* 使他的条件盈利达到极大化。这个定义与正则型博弈中 Nash 均衡的基本思想是一样的，因此在检验某 s^* 是否 Nash 均衡时，我们总是在“固定”局中人 i 的对手的策略条件下检验 i 是否愿意偏离 s 路径，但应注意到这样做并没有要求局中人的策略必须“同时地”选取。我们将从以后的若干例子中有所体会。

定义了展开型博弈的纯策略剖面 s^* 以及纯策略 Nash 均衡，接下来自然想到应该定义展开型博弈的混合策略以及相应的混合策略均衡了。

正则型博弈中的混合策略实际上是在纯策略空间上的概率分布，因此在我们已经定义了展开型博弈中局中人 i 的纯策略空间 S_i 之后，只要在 S_i 确定的每一个概率分布将都是局中人 i 的混合策略。用 Luce 与 Raiff 的比喻来说，局中人 i 的每一个纯策略 s_i 相当于一本指导说明书，书的每一页，表示到了一个特定的信息集 h_i，在该页上告诉了局中人 i 在该信息集上如何行动。许许多多的 s_i 相当于有许许多多的说明书，S_i 表示这些说明书的全体，它们被安放在一个大书架上。混合策略就相当于局中人 i 以一定的概率分布从书架上随机地抽取一本说明书。因此只要将 S_i“整理”好。混合策略也随之可以明确地定义。

有一点似乎可以设想，倘若有这样的一本说明书，局中人 i 翻

到每一页，也就是他到达具有某信息集的一个决策结，这一页上列举了基于该信息集 h_i 的所有行动 $A(h_i)$，并告诉他如何随机地从 $A(h_i)$ 中选取一个行动，那么读这本说明书就几乎相当于从那么大的书架上随机地取书。这就是我们这里要介绍的行为策略(behavior strategies)。现在我们将在 $A(h_i)$ 上确定概率分布，就是说，对 $A(h_i)$ 中的每个行动赋予一定的概率。这样可以得到许许多多不同的概率分布，$A(h_i)$ 上所有概率分布的全体不妨记为 $\Delta(A(h_i))$。对应于 H_i 中的每一个 h_i，均有 $\Delta(A(h_i))$ 与之对应。纯策略空间 S_i 既然定义作 $A(h_i)(h_i \in H_i)$ 的笛卡儿乘积空间。那么相应概率分布的乘积空间 $\times_{h_i \in H_i}\Delta(A(h_i))$ 可以看作 S_i 的大幅度扩展，因为 $A(h_i)$ 中每一个的行动都可以视作 $\Delta(A(h_i))$ 内的一个退化分布。局中人 i 的行为策略 b_i 将定义为 $X_{h_i \in H_i}\Delta(A(h_i))$ 内的一个元素。N 个局中人的行为策略剖面可记作

$$b=(b_1,\cdots,b_n)$$

本身就产生结局上的概率分布，因此对每一个局中人 i 产生一个期望盈利。行为策略的 Nash 均衡是这样的一个策略剖面，它使得没有一个局中人可以通过使用不同的行为策略而增加自身的盈利或效用。

2. 展开型博弈的策略型表示

先考虑一个简单的展开型博弈(如图 4.5 所示)。

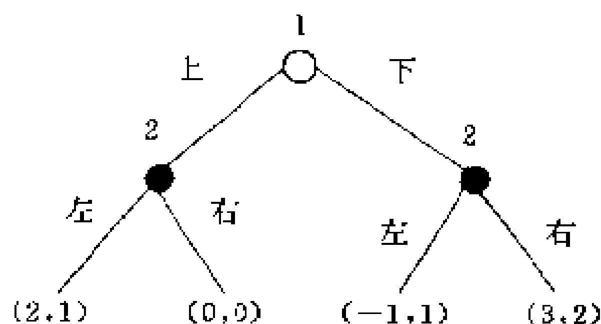


图4.5

根据我们迄今为止累积的有关知识易知，图 4.5 所示的博弈中，局中人 1 有两个纯策略，$S_1=\{(上),(下)\}$，局中人 2 有四个纯策略，$S_2=\{(左,左),(左,右),(右,左),(右,右)\}$。依据两个局中人的纯策略，我们不难构成纯策略剖面如下：

(1){上,(左,左)}，路径为(1,2(左),(2,1))

注：属于局中人 2 的信息集有左与右 2 个，我们将它们分别记作 2(左)与 2(右)，终点结干脆以盈利向量表示，顺便指出了该剖面的各局中人的盈利，以下同。

(2){上,(左,右)}，路径为(1,2(左),(2,1))

(3){上,(右,左)}，路径为(1,2(左),(0,0))

(4){上,(右,右)}，路径为(1,2(左),(0,0))

(5){下,(左,左)}，路径为(1,2(右),(−1,1))

(6){下,(左,右)}，路径为(1,2(右),(3,2))

(7){下,(右,左)}，路径为(1,2(右),(−1,1))

(8){下,(右,右)}，路径为(1,2(右),(3,2))

从上述罗列中可发现一个有趣的事实：不同的纯策略剖面可以有相同的路径与结局，例如(1)与(2)、(3)与(4)等。上面 8 种情况很容易列成如图 4.6 所示的盈利矩阵。

		局中人 2			
		(左,左)	(左,右)	(右,左)	(右,右)
局中人 1	上	2,1	2,1	0,0	0,0
	下	−1,−1	3,2	−1,1	3,2

图 4.6

读者不难发现，图 4.6 相当于图 4.5 展开型博弈的策略型表示。图 4.6 蕴含着局中人 1 与 2 在行动之前均预先作出一个全面的应急计划，比如(左,左)就是局中人 2 周到地考虑到局中人 1 取“上”或“下”时他的一个可能应付，再加上(左,右)，(右,左)及(右,右)则使整个博弈中每一种应该考虑的可能情况全部列入计划。而

本质上同样的博弈表示成图 4.5 那种展开型式，局中人 2 在决定自己究竟取左与右两个行动中的哪一个时要等到信息集 h_2 已达到，看看 h_2=“上”还是 h_2=“下”再下手。

现在考察如图 4.7 和图 4.8 所示的两个展开型博弈。

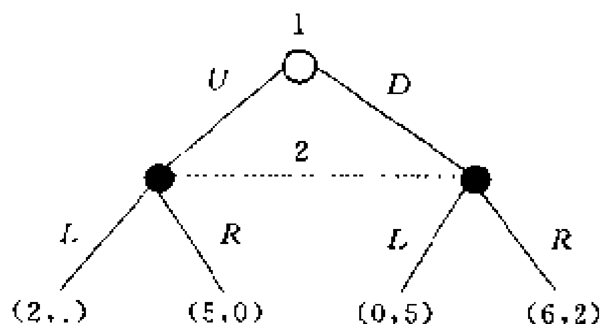


图4.7

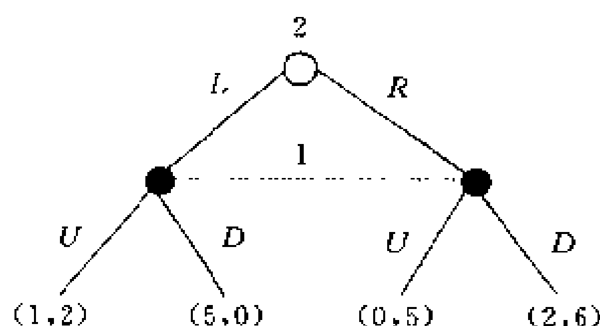


图4.8

注：图 4.8 所示的博弈中，盈利向量的第一个元素是局中人 2 的盈利。

很容易将这两个展开型博弈表示成如图 4.9 所示的同一的策略型。

局中人 2

		L	R
局中人 1	U	2,1	5,0
	D	0,5	6,2

图 4.9

两个型式不同的展开型表示为同一策略型，在这里是不足为怪的。因为这是一个双方同时行动的静态博弈。不存在先行者，在

将这个博弈展开成博弈树时，任意一个局中人都可以放在初始结，只不过另一个局中人将处于一个不完美信息集。

其实，任何有限个局中人的展开型博弈(尤其是二人博弈)，不管展开层次有多少(但为有限次)，根据各个局中人的纯策略空间的建立，我们总可以用策略型来表示。问题在于，纯策略空间构造方式清晰地告诉我们，随着层次的展开，纯策略的个数将迅速地增多，这给策略型表示平添了不少麻烦。例如，如图 4.10 所示的一个简单的展开型博弈：

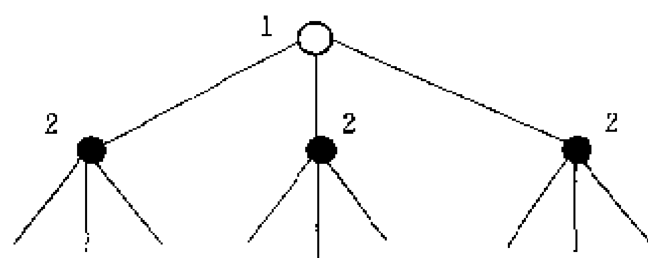


图4.10

非常清楚，博弈树的路径有 9 条。但从纯策略个数来看，局中人 1 有 3 个，局中人 2 有 $3\times3\times3=27$ 个。如将它表示为策略型，盈利矩阵为 3×27 阶，似乎还不及图 4.10 一目了然。麻烦的原因之一也许在于策略空间太大了。

纯策略个数太多，只要合理，也就无可指责。但是，有时我们会发现某些纯策略在不管对手如何行动的情况下产生相同的效果，例如见图 4.11。

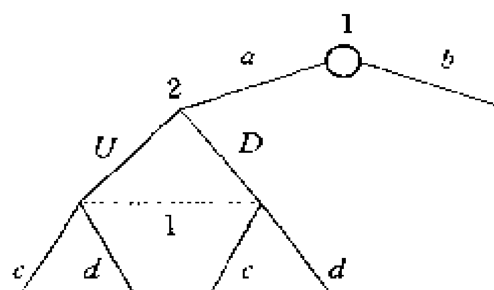


图4.11

局中人 1 有两个信息集 h_1^0、h_1^1，其中 $A(h_1^0)=\{a,b\}$，$A(h_1^1)=\{c,d\}$，因此局中人 1 的纯策略空间为

$$S_1=\{A(h_1^0),A(h_1^1)\}=\{(a,c),(a,d),(b,c),(b,d)\} \quad (4.4)$$

显然，如果局中人 1 取行动 b（注意：不能说局中人 1 取策略 b，因为他的纯策略是整个博弈中他的计划，因此策略只能来自于(4.4)式中 4 个元素），相应的纯策略必为 (b,c) 与 (b,d)，毫无疑问，这两个纯策略中的第二个信息集不可能达到，因此无论局中人 2 采取行动 U 或者 D，(b,c) 与 (b,d) 的结局完全是一样的，鉴于该事实，人们完全有理由将 (b,c) 与 (b,d) 两个纯策略"合二为一"。在下面，我们将首先引入等价纯策略的概念，为了定义的普遍性，总是认为在初始结的"自然"所选择的行动空间具有概率分布，故在所有结局上相应地具有概率分布。

定义 4.1 如果展开型博弈中局中人 i 的两个纯策略 s_i 与 s_i' 对于其对手的所有纯策略产生博弈结局上的同一概率分布，则称 s_i 与 s_i' 是等价的。

于是，图 4.11 中局中人 1 的纯策略 (b,c) 与 (b,d) 是等价的。

定义 4.2 展开型博弈的简化策略型(reduced strategic form)可以通过在每个等价纯策略类中只留下一个而去除其余等价纯策略的方法得到。

按照定义 4.2，我们只在等价纯策略类中选定代表从而构成简化策略型，简化策略型中的纯策略再也不可能等价。在展开型博弈的简化策略型中，局中人 i 的每个纯策略导致不一样的结果。我们前面所说的基于纯策略空间定义混合策略，如果改为在简化策略型的纯策略空间上的概率分布也许在定义方面更合理。今后我们如果提及混合策略，其定义必建立在简化策略型上。

即使这样定义的(似乎更合理的)混合策略与我们前一节介绍的行为策略在定义上仍有所不同,行为策略是在 $A(h_i)$ 上随机化,而混合策略则在 $A(h_i)$ 的乘积空间(简化)上随机化。但是它们的共同点是"随机化",可以想象它们之间肯定有关系,在经济文献中的绝大部分博弈,由于博弈的"完美回忆",混合策略与行为策略是等价的。我们将在下小节介绍这个等价性。

3. 完美回忆中混合策略与行为策略的等价性

经济文献中几乎所有博弈是完美回忆博弈(prefect recall),所谓完美回忆是指博弈过程中,没有一个局中人在任何时候忘记他曾经知道的信息,且所有局中人都知道他们先前已经采取过的行动。例如,讨价还价模型就是完美回忆博弈。因此研究完美记忆下混合策略与行为策略之间的等价性具有实践意义,它为我们今后的经济博弈模型研究带来了方便。

这里所研究的等价性还是前一小节提及的"等价性";两个策略称作等价的,如果对于其他局中人所有策略所形成的各种结局上它们导致相同的概率分布。我们的讨论先从由任何一个混合策略出发产生唯一的行为策略开始,由于行为策略是在每一个 $A(h_i)$ 上随机选取行动,因此我们考虑的混合策略也只能建立在所有 $A(h_i)$ 的乘积空间的基础上。也就是说,令混合策略 σ_i 基于纯策略空间 S_i(而不是简化策略型)之上。对于局中人 i 的每一个历史 h_i,该局中人的纯策略 s_i 与其他局中人的纯策略 s_{-i} 构成纯策略剖面 s,有些 s 的路径包含 h_i,而有些 s 的路径不包含信息集 h_i,对于局中人 i,那些使得 $s=(s_i,s_{-i})$ 包含 h_i 的纯策略 s_i 的集合记作 $R_i(h_i)$,也就是说,如果 $s_i\in R_i(h_i)$,那么存在局中人 i 的对手的纯策略剖面 s_{-i} 到达信息集 h_i。如果混合策略 σ_i 对于 $s_i\in R_i(h_i)$ 赋予正概率,那么在信息集 h_i 的行动集合 $A(h_i)$ 中的每一个行动 a_i(即 $a_i\in A(h_i)$),我们定义概率 b_i 如下:

$$b_i(a_i|h_i)=\sum_{\{s_i\in R_i(h_i)且s_i(h_i)=a_i\}}\sigma_i(s_i)/\sum_{\{s_i\in R_i(h_i)\}}\sigma_i(s_i) \tag{4.5}$$

即利用混合策略定义 $R_i(h_i)$ 中恰好取行动 a_i 的那些 s_i 的条件概率。如果对所有的 $s_i\in R_i(h_i)$，混合策略 σ_i 均赋予零概率，此时(4.5)式的条件概率定义无意义，我们令

$$b_i(a_i|h_i)=\sum_{\{s_i(h_i)=a_i\}}\sigma_i(s_i) \tag{4.6}$$

非负数 $b_i(a_i|h_i)$ 显然满足

$$\sum_{a_i\in A(h_i)}b_i(a_i|h_i)=1 \tag{4.7}$$

这是因为每一个 s_i 确定了局中人 i 在信息集 h_i 的一个行动，于是我们利用了混合策略 σ_i 为局中人 i 在信息集 h_i 的行动集 $A(h_i)$ 建立了一个随机选取行动的概率分布。或者说，根据任意一个混合策略 σ_i，产生了唯一的行为策略 b_i。从概率论的知识，我们使用的记号 $b_i(a_i|h_i)$ 中，当 $a_i\in A(h_i)$ 时，条件 h_i 其实是多余的，可是“条件”本身有助于强调 a_i 是信息集 h_i 上一个可行的行动。

我们用一个例子来详细阐述由混合策略构造行为策略。

例 4.2 展开型博弈如图 4.12 所示。

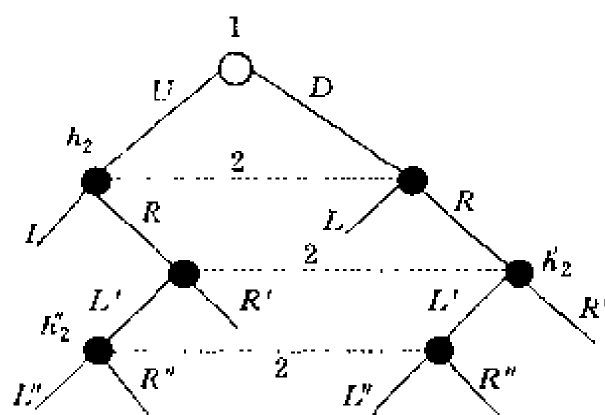


图4.12

图 4.12 描述这样的博弈过程：局中人 1 的行动为 U 与 D 两种，局中人 2 在不知道局中人 1 的行动情况下采取 1 次或 2 次或 3 次行动。因此局中人 2 有三个信息集 h_2、h'_2、h''_2，各具有行动集 A

$(h_2)=\{L,R\}$，$A(h_2')=\{L',R'\}$、$A(h_2'')=\{L'',R''\}$。依定义，局中人2的纯策略应有 $2\times2\times2=8$ 种。现考虑局中人2的混合策略 σ_2 仅对两个纯策略：$s_2^1=(L,L',R'')$ 与 $s_2^2=(R,R',L'')$ 各赋予1/2概率。由此产生的(等价)行为策略为：

$$b_2(L|h_2)=b_2(R|h_2)=(1/2)/(1/2+1/2)=1/2$$

$$b_2(L'|h_2')=0/(1/2)=0$$

$$b_2(R'|h_2')=(1/2)/(1/2)=1$$

$$b_2(L''|h_2'')=\sigma_2(s_2^2)=1/2$$

$$b_2(R''|h_2'')=\sigma_2(s_2^1)=1/2$$

注意，上述行为策略计算中最后两次使用了(4.6)式。

从构造方法即(4.5)式、(4.6)式可知，每一个混合策略一定产生唯一的行为策略。有趣的问题在于不同的混合策略是否会产生同样的行为策略呢？回答是肯定的。请看下面的例子：

例4.3 考虑二人动态博弈，如图4.13所示。

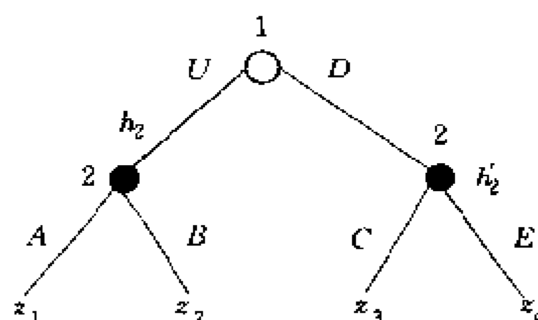


图4.13

局中人2有两个信息集 h_2 与 h_2'，各自行动集合为 $A(h_2)=\{A,B\}$，$A(h_2')=\{C,E\}$。故纯策略空间 $S_2=\{(A,C),(A,E),(B,C),(B,E)\}$。现有两个混合策略：$\sigma_2^1=(1/4,1/4,1/4,1/4)$，即对四个纯策略赋予相等的概率；$\sigma_2^2=(1/2,0,0,1/2)$，对 (A,C) 与 (B,E) 各赋予概率1/2。先计算由 σ_2^1 产生的行为策略：

$$b_2(A|h_2)=b_2(B|h_2)$$
$$=(1/4+1/4)/(1/4+1/4+1/4+1/4)=1/2$$

$$b_2(C|h_2')=b_2(E|h_2')$$
$$=(1/4+1/4)/(1/4+1/4+1/4+1/4)=1/2$$

再计算由 σ_2^2 产生的行为策略：

$$b_2'(A|h_2)=b_2'(B|h_2)=(1/2)/(1/2+1/2)=1/2$$
$$b_2'(C|h_2')=b_2'(E|h_2')=(1/2)/(1/2+1/2)=1/2$$

显然，这两个行为策略是相同的。

设想局中人 1 采取混合策略 σ_1：以概率 p 取 U。以概率 $1-p$ 取行动 D，那么不管局中人 2 取什么样的混合策略 σ_2^*，如果 b_2^* 是由 σ_2^* 产生的行为策略，各终点结 $z_j(j=1,2,3,4)$ 到达的概率计算为（以下 P_1、P_2 分别表示局中人 1、2 的行动概率）：

$$\left.\begin{aligned}P(z_1)&=P_1(U)\cdot P_2(A|U)=p\cdot b_2^*(A|h_2)\\P(z_2)&=P_1(U)\cdot P_2(B|U)=p\cdot b_2^*(B|h_2)\\P(z_3)&=P_1(D)\cdot P_2(C|D)=(1-p)\cdot b_2^*(C|h_2')\\P(z_4)&=P_1(D)\cdot P_2(E|D)=(1-p)\cdot b_2^*(E|h_2')\end{aligned}\right\}\quad(4.8)$$

由于 σ_2^1 与 σ_2^2 产生同样的行为策略 b_2，由(4.8)式可知，σ_2^1、σ_2^2 与行为策略 b_2 导致终点结上相同的概率分布。因此，σ_2^1、σ_2^2、b_2 是等价的。

图 4.13 所表示的展开型博弈是具完美回忆的，因此混合策略等价于它所产生的唯一的行为策略，而且每一个行为策略等价于产生它的每一个混合策略。这就是 1953 年由 Kuhn 得出的结论：

定理 4.1 （Kuhn 1953）在完美回忆博弈中，混合策略与行为策略是等价的。

我们关心的是，如果博弈不具有完美回忆，混合策略与行为策略之间的等价关系是否会发生问题。让我们考虑如图 4.14 所示的博弈。

图 4.14 描述了一个不具完美回忆的博弈。局中人 1 在信息

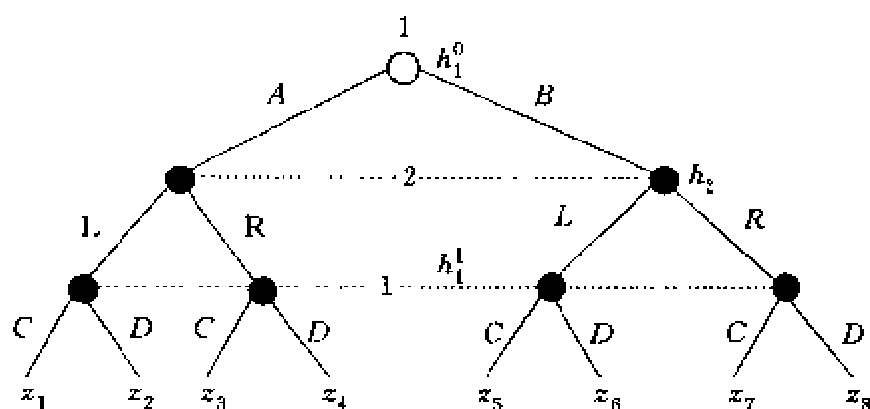


图4.14

集 h_1^1 时显然患了健忘症，他忘记了自己从初始结出发时到底采取了行动 A 还是行动 B。局中人 1 有两个信息集：h_1^0（其实 $h_1^0=\varnothing$ 为空集）与 h_1^1，因此他的纯策略有四个：

$$s_1^1=(A,C),\ s_1^2=(A,D),\ s_1^3=(B,C),\ s_1^4=(B,D)$$

考虑混合策略 $\sigma_1=(1/2,0,0,1/2)$，计算由 σ_1 产生的行为策略：

$$b_1(A|h_1^0)=b_1(B|h_1^0)=1/2/(1/2+1/2)=1/2$$

$$b_1(C|h_1^1)=b_1(D|h_1^1)=1/2/(1/2+1/2)=1/2$$

即 $b_1=\{(1/2,1/2),(1/2,1/2)\}$，局中人 1 在信息集 h_1^0 与 h_1^1 均以概率 1/2 取行动集中的一个元素。我们可以发现 b_1 与 σ_1 不是等价的。注意到局中人 2 只有两个纯策略 L 与 R。若局中人 2 取 $s_2=L$，或理解为局中人 2 以概率 1 取 L。取定 $s_2=L$ 时，我们可以计算在 (σ_1,L) 下到达终点结的概率：

$$P(z_1)=P_1(A,C)\times P_2(L|A)=\sigma_1(A,C)\times 1=1/2$$

$$P(z_6)=P_1(B,D)\times P_2(L|B)=\sigma_1(B,D)\times 1=1/2$$

显然其他终点结的到达概率为 0。再考察策略剖面 (b_1,L)，z_1 通过路径 (A,L,C) 到达，由于 $b_1(A|h_1^0)=b_1(C|h_1^1)=1/2$，行为策略描述了在不同信息集上随机取行动的独立性，因此 $P(z_1)=1/2\times 1/2=1/4$。同样理由，通过 (A,L,D) 到达 z_2 的概率为 1/4；通过 (B,L,C) 到达 z_5 的概率为 1/4；通过 (B,L,D) 到达 z_6 的概率为

1/4。可见 σ_1 与由 σ_1 产生的 b_1 不是等价的。其原因在于局中人 1 使用混合策略 σ_1 时，若取 A 则然后必取 C，若取 B 则然后必取 D（因为 (A,C) 与 (B,D) 各取 1/2 概率），使得在使用混合策略 σ_1 时局中人 1 在信息集 h_1^0 与 h_1^1 所取行动具有一定的相关性。而对于行为策略，由于局中人 1 的健忘性，他在信息集 h_1^1 选择行动时已经忘记了自己在 h_1^0 时究竟是取了 A 还是取了 B，于是 b_1 告诉他在 h_1^1 时等可能地选择 C 或 D，于是原先 σ_1 下"A 必 C"、"B 必 D"的相关性消失了，因此这两种策略不可能等价。倘若将图 4.14 中信息集 h_1^1 三段虚线中间的那段去除，局中人 1"恢复"了记忆，此时由 σ_1 产生的 b_1 与 σ_1 根据 Kuhn 定理必然等价。

由于本书主要介绍与研究经济博弈模型，因此我们今后遇到的模型是完美回忆的，混合策略与行为策略之间的等价性为我们今后的叙述与记号带来了方便。例如我们将以 $\sigma_i(a_i|h_i)$ 表示局中人 i 在信息集 h_i 取 a_i 的概率，而不另起用记号 b_i。

4. 累次严优与 Nash 均衡

展开型博弈可以表示成策略型式，如我们所指出过的那样，这种方法有时反而显得复杂，其实用性受到一定怀疑。但是从理论角度却极有意义。例如，如果展开型是有限步，那么相应的策略型必也是有限的，有限正则型博弈至少存在 Nash 均衡（可能是混合策略），因而有限展开型博弈至少存在 Nash 均衡也就顺乎自然。想当然地，累次严优方法也可以延伸到展开型博弈中，可是在大多数展开型博弈中这个概念有点勉强。问题出在，在给定对手行动时，不能达到的信息集上局中人不能严格地喜欢某个行动甚于另一个行动。展开型博弈的结局维数一般地显然低于利用局中人的全部纯策略构成的策略型的维数，我们从前面的策略型表示的例子中也可以看出，在策略型表示中会出现二个策略剖面导致同一终点结的现象，因此在比较策略的优劣时遇到了不必要的麻烦，至少原先可能严劣的策略变为弱劣策略或者其他什么情况。这个现象使

我们意识到基于普通策略型盈利或效用的定理也许并不适合于展开型博弈。

像累次严优解关于正则型博弈是博弈的合理预测(若存在累次严优解的话),展开型博弈的一个可能的合理推断方法就是我们下一节将要介绍的后退归纳法。在这小段里,我们利用动态博弈的后退归纳思想证明一条重要的结论:

定理 4.2 (Zermelo 1913,Kuhn 1953)完美信息的有限博弈具有一个纯策略 Nash 均衡。

首先注意到,完美信息博弈的所有信息集都是单一集,即每个信息集中只有一个结。既然博弈是有限的,必定存在倒数第二个结的集合,这些结中每一个的后续结是终点结。确定在倒数第二个结中每一个均可行动的局中人选择一个行动以使他在将到达的终点结具有最高盈利(或效用)。例如,图 4.15 就是一个虚拟的完美信息博弈。

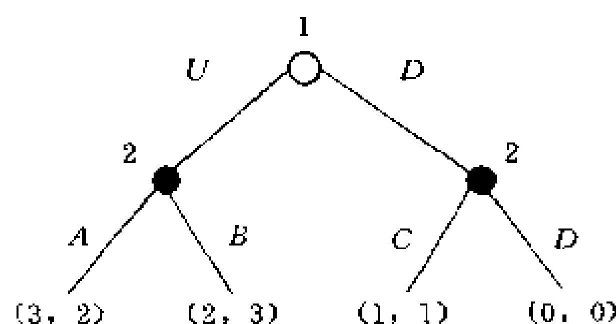


图4.15

该博弈树有两个倒数第二个结,且均为局中人 2 的决策结。在四种可能结局中,局中人 2 取 B,则获最高盈利 3。也可能存在这样的现象,对该局中人来说,有两个或两个以上的结局都使他获得同样的最高盈利,称为存在"结"(tie)现象(注:此结与博弈树中的结(node)有着本质差异),那么可以在其中任意选择一个行动。现在在给定处于倒数第二结的局中人采取我们刚才已确定的行动

前提下，确定在其直接后续结是倒数第二结的那些结上的每一个局中人选择一个在可行后续结上盈利为极大的行动。这样我们可以通过博弈树一步一步地往后退，在每个结上为该行动的局中人选定行动。事实上，我们已经为每一个局中人确定了一个策略。这些策略形成了一个 Nash 均衡(注：这里没有证明它是 Nash 均衡，但在下一章中我们介绍了“一步偏离准则”之后读者不难验证这一条)。以上文字叙述也许使有的读者感到不够直观，我们在下一章介绍后退归纳法时将以实例进行详尽阐述。不过请读者注意，这里叙述的办法(数学上称为 Zermelo 算法)与后退归纳有所区别，Zermelo 算法在沿博弈树后退过程在每一个结为每一个局中人确定其最大可能盈利的行动，实际上 Zermelo 算法是后退归纳法的多人($n>2$)博弈推广。

如果定理 4.2 条件不满足的话，Zermelo 算法就难以有所作为。例如，如果将有限博弈改为无限博弈，无非是要么具有一个单一结，在它后面有无限多个后续结(譬如博弈具有连续统的行动空间)，要么博弈具有一条由无限多个结组成的路径。在后一种无限博弈情况，不存在倒数第二结，故无法谈论 Zermelo 算法；在第一种情况，如果对盈利函数没有进一步限制的话则不存在最佳选择，简单地说，即使行动空间是[0,1]这样具连续统势的闭区间。并不是说[0,1]上的函数一定存在极大值，必须要对盈利函数加上适当的条件。其次，如果定理条件中的完美信息改为不完美信息的话，设想倒数第二结相应的信息集不完美，处于该信息集的局中人无法知道自己处于哪个决策结，纵然信息是完全的(即他对各结局的盈利向量一清二楚)，他还是不能为自己确定一个最佳行动。最简单的例子是猜谜博弈，由于两个局中人同时行动，因此信息不完美，绘制博弈树如图 4.16，图中枝旁小圆圈中的数字表示行动中伸出的手指数。

倒数第二结是局中人 2 的决策结，共有 2 个。但它们构成局

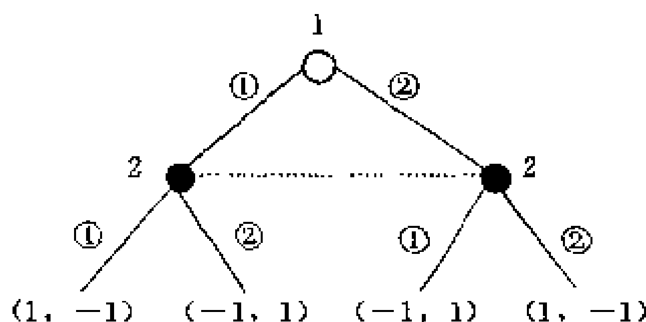


图4.16

中人2的不完美信息集,局中人2的可能最高盈利为1,可是由于他不知道局中人1伸出几根手指,因此他确定不了自己伸出多少根手指为最佳选择。事实上,我们已经知道,猜谜博弈不存在纯策略Nash均衡。

§4.3 Stackelberg博弈与后退归纳法

迄今为止我们只讨论了选择为离散的策略与行动。展开型博弈也存在行动空间为连续统的情况。其实在正则型博弈中,我们已经遇到过连续的纯策略空间的博弈模型,著名的例子就是Cournot竞争,在那里,两个公司考虑的策略是有关产量多少的选择。现在我们假定博弈存在一个过程,其中一个局中人为先行者,不妨设为公司1,首先选择产量水平q_1,为了讨论与计算的方便,同时也不影响到问题的实质,不妨假设生产无需成本。公司2在观察到公司1的q_1后,选择自己的产量水平q_2。q_1与q_2来自空间$Q_1=Q_2=[0,\infty)$。现令市场需求函数呈线性状态:

$$p(q_1+q_2)=12-(q_1+q_2) \tag{4.9}$$

公司1与公司2的盈利函数为

$$u_i(q_1,q_2)=[12-(q_1+q_2)]q_i \qquad (i=1,2) \tag{4.10}$$

由于它们的策略选择存在先后顺序,因此博弈是动态的,不能像原先Cournot竞争模型那样预测结局。然而由于可供选择的行动空

间是连续区间，无法用博弈树表述，但博弈进行过程十分清楚：

公司 1 先"简单地"选择 q_1。

公司 2 在观察到 q_1 后，确定自己的策略 s_2，s_2 是由 Q_1 到 Q_2 的映射，对于每一个观察到的 q_1，局中人 2 必定选择 q_2 以使 $u_2(q_1, q_2)$ 达到极大。根据一阶条件易得 $r_2(q_1)=6-q_1/2$。Nash 均衡则要求公司 1 在给定 $s_2=r_2$ 时极大化 $u_1(q_1,q_2)$。尽管他观察不到 r_2，理性的他一定知道这个 r_2 是在公司 2 观察到自己的 q_1 之后的最佳反应，因此 $r_2(q_1)=6-q_1/2$ 也是公司 1 可以预测到的，此时公司 1 的盈利变成为

$$u_1(q_1,q_2)=[12-(q_1+q_2)]q_1=(6-q_1/2)q_1 \qquad (4.11)$$

再由一阶条件知 $q_1=6$，于是 $q_2=6-q_1/2=3$。相应地，$u_1=18$，$u_2=9$，产量水平(6,3)被称为 Stackelberg 均衡。由于市场需求的牵制，$r_2(q_.)$ 通常是 q_1 的递减函数，正应了"先下手为强"这句话，公司 1 通过先行并增加自己的产量而减少了对手的产量，其结果是，在 Stackelberg 均衡中，公司 1 达到产量与盈利均高于 Cournot 均衡中他的产量与盈利；相反地，公司 2 的产量与盈利却低于他在 Cournot 均衡中的水平。这是动态博弈中所谓先行优势。

Stackelberg 均衡的推算当然借用了 Nash 均衡的基本思想，因而自然地用来作为博弈的预测结果。在这种"先行后动"的竞争模型中是否只存在唯一的 Nash 均衡呢？不妨设想，作为理性的公司 2 或许可以认识如下事实：尽管公司 1 最先选择 q_1、在完全信息下进行动态博弈，公司 1 在选择 q_1 时必定应该考虑到公司 2 的反应。因此尽管行动有前有后，公司 2 基于 q_1 去选择 q_2，公司 1 理应考虑到 q_2 将随 q_1 而变，从而在选择 q_1 时就将这个因素考虑进去。于是似乎问题又被"牵回"到"两人同时行动模型"。此时 Nash 均衡为 Cournot 竞争中的(4,4)，相应盈利向量为(16,16)。这个结果对公司 2 来说要好于他在 Stackelberg 均衡时的处境，基于极大化自身盈利的思想，不管公司 1 先前如何行动，公司 2"以不变应万

变"，取 $q_2^c=4$，从而"迫使"公司 1 在此"威胁"下取 $q_1^c=4$。公司 2 不愿偏离 q_2^c 以免蹈 Stackelberg 均衡的覆辙，公司 1 接受威胁不愿偏离 q_1^c 以免使自己处境更糟，谁也不愿偏离的策略剖面 $(q_1^c,q_2^c)=(4,4)$ 当然也是 Stackelberg 博弈的 Nash 均衡。

如同完全信息静态博弈中的多重 Nash 均衡一样，在动态的 Stackelberg 模型中我们又面临了两个均衡，哪一个作为博弈的预测结果更合理一些呢？看来公司 2 的威胁的可信度是个关键。只要公司 1 认为公司 2 是个理性的竞争对手，他就认为威胁是不可信的，毕竟他处于先行地位，他向公司 2"出示"自己的选择 $q_1=6$，这种场合公司 2 再坚持取 $q_2^c=4$ 的盈利肯定低于他取 $q_2=3$ 的盈利。可见 Nash 均衡(4,4)依赖于一个空头威胁，因此在预测博弈结果时舍弃 $(q_1^c,q_2^c)=(4,4)$ 而采用 Stackelberg 均衡。当然，如第二部分开头所叙述，公司 1 也可能怀疑威胁取 $q_2^c=4$ 的公司 2 不是理性的，我们以后将会在建模时让这类怀疑转为不完全信息——公司 1 无法清楚公司 2 的盈利函数。

在 Stackelberg 博弈中，容易知道公司 2 应该如何行动，因为一旦 q_1 被确定之后实际上公司 2 面临一个简单的决策问题，于是我们对每一个给定的 q_1 求解公司 2 的最佳选择，再回过头去寻找公司 1 关于 q_1 的最佳选择。这种逻辑推理方法就是所谓后退归纳法(backward induction)。在对博弈出现多重 Nasb 均衡需要"精炼"舍弃较不合理者时，它是有效方法之一。

后退归纳法包括以下几个步骤：

(1)从博弈的终点结出发。追踪每一个终点结到紧接它前面的结(即 Zermelo 算法中的倒数第二结)。它们是某些局中人的决策结，这些决策结分为"平凡的"、"基本的"与"复杂的"。如果决策结的每一根"枝"正好导向一个终点结则称该决策结是基本的；只拥有一根枝的基本决策结称为平凡的。非基本结则称为复杂结。如果到达一个平凡的决策结，那么继续向博弈树上方移动直到你到

达一个要么是复杂的要么是非平凡的基本决策结为止，或者直至你不能继续上行为止。

(2)在步骤(1)中到达的每一个基本决策结上，通过对由该决策结出发到达的每一个终点结上局中人得到的盈利求最佳行动(最佳行动可能不止一个，也就是存在"结"(tie)，在 Zermelo 与 Kuhn 的定理中，由于只要求证明存在一个纯策略 Nash 均衡，因此我们只需要从中任意地选取一个最佳行动。但在这里，我们追求的是合理的 Nash 均衡，这样就要求我们去确定是否存在"社会习俗或惯例"，使得其中一个选择比其余的更合适一些，倘若没有任何惯例可以帮助，那么我们必须承认在这个结上局中人的习性是无法预测的)。注意在基本决策结(设为 A)与终点结(设为 B)之间的每一条路径只从 A 的某一枝出发。导致局中人获得最高盈利的枝就是在该结所作的最佳行动。

(3)将在步骤(2)中检验过的每一个基本决策结所引起的所有非优枝删去。使得这些基本决策结的每一个都成为平凡的。

(4)现在有了一个新的博弈树，它比原来的那个当然要简单些。如果在步骤(1)已经到达树根，则到此结束。

(5)如果尚未到达树根，那么回到第(1)步。对新博弈树重新开始(1)至(4)过程。以这种方式，一步一步地走向树根。

(6)对每一个局中人，将该局中人在每一个决策结上的最优决策一起收集起来，这个决策集合构成了博弈中局中人的最佳策略。

以上面 6 个步骤综述的后退归纳法显然是针对有限完美信息博弈的。注意到它与 Zermelo 算法"原则上"是一回事。如果信息不完美，如同对定理 4.2 所作的解释那样，后退归纳实际上无法奏效。当然在无限博弈时也会遇到麻烦(但在以后章节中会阐述一些特殊情况)。当行动空间具有连续统时，由于无法用博弈树表达，上面的步骤一般不适用，但像 Stackelberg 博弈那样的后退归纳，我们在前面已略作介绍。下面以简单的例子来对后退归纳有直观了

解。

例 4.4 展开型博弈如图 4.17 所示。

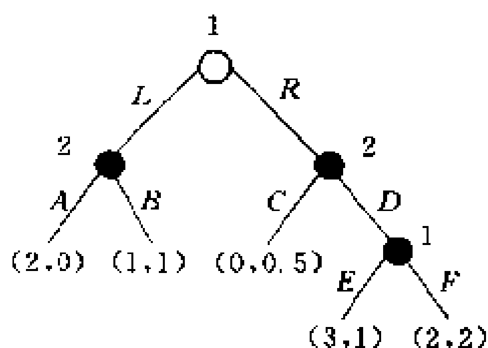


图4.17

根据第一轮删弃得到新博弈树如图 4.18 所示。

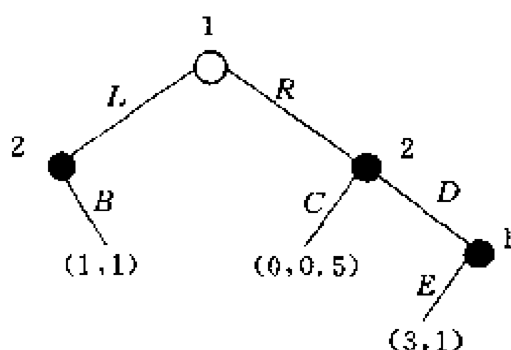


图4.18

经过第二轮删弃又得到更新的博弈树如图 4.19 所示。

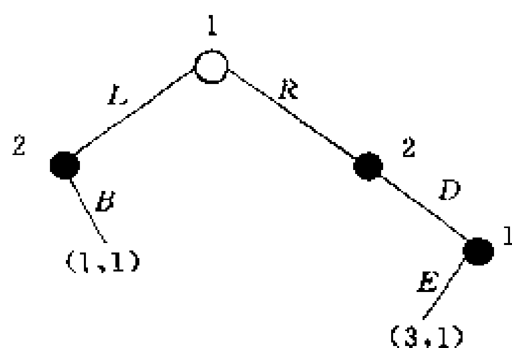


图4.19

因此该博弈的后退归纳解为$\{(R,E),(B,D)\}$。

按照步骤(6)的说法，后退归纳解实际上为每一个局中人选择

了最优策略，因此它是博弈的 Nash 均衡，我们有如下结论，它是博弈论中一个重要命题。

定理 4.3 在一个具有完美信息的有限博弈中，使用后退归纳选择的策略剖面总是 Nash 均衡。

§4.4 子博弈和子博弈完美均衡

Stackelberg 博弈中的 Nash 均衡(4,4)由于局中人 2 的不可信威胁而被我们从预测中剔除。现在，博弈论专家几乎普遍地接受这样的事实：一个掺合着不可信威胁的 Nash 均衡对于人类习性将是一个差劲的预测。后退归纳法是剔除这种不合理 Nash 均衡的有力工具之一，但它的应用具有局限性。Selten 发明了子博弈(subgame)概念和称之为子博弈完美均衡(Subgame Perfect Equilihrium)解的概念，以使“合理的”Nash 均衡与“不合理的”Nash 均衡分离。正因为这个贡献，他与 Nash、Harsanyi 一起获得 1994 年诺贝尔经济学奖。

Selten 首先提出在一般展开型博弈中，某些 Nash 均衡比其他均衡更合理。著名的例子是如图 4.20 所示的博弈。

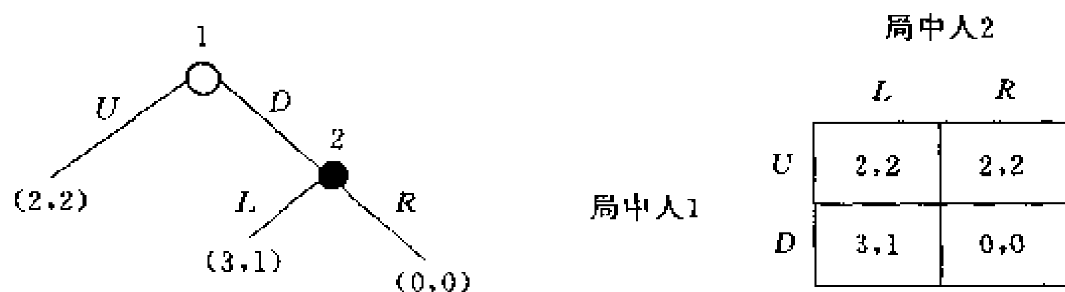


图4.20

图 4.20 的左边是展开型博弈树，右边则是它的策略型表示。从右图通过划线法易知该博弈有两个纯策略 Nash 均衡：(U,R)

与(D,L)。当时，Selton 就指出(U,R)是可疑的，因为它依赖于局中人 2 取 R 这一"空头威胁"，局中人 1 不相信这个空头威胁而取行动 D 时，局中人 2 的信息集达到，此时理性的他也不会希望实施 R 而只得取 L 以使自己可以获盈利 1。现在我们对 Selten 的观点更是深信不疑，因为我们很容易从图 4.20 的左图求得它的后退归纳解是(D,L)。倘若我们将图 4.20 的博弈树稍微延伸一点，见图 4.21。

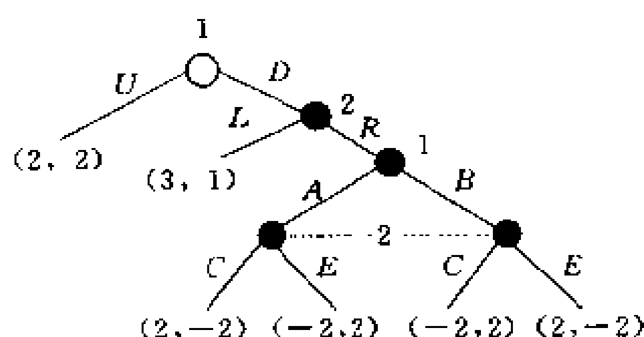


图4.21

在延伸部分，局中人 2 处于不完美信息集，他不清楚局中人 1 取了 A 还是取了 B，在这种情况下，行动 C 与 E 之间不存在孰优孰劣，后退归纳法无法应用。但是我们仔细地观察一下，若以局中人 1 的第二个决策结作为初始结的话，那么延伸部分恰为完全信息静态博弈中的猜谜游戏。已知猜谜博弈具有唯一的混合策略 Nash 均衡解$\{(1/2,1/2),(1/2,1/2)\}$，它使两个局中人均具有平均盈利 O。以延伸部分作为整体当作一个"终点结"来处理，其盈利以延伸部分的平均盈利$(0,0)$代之，从逻辑上来说有其合理性，从形式上看又回到了图 4.20 的情况，又可应用后退归纳法了。这是因为若局中人 1 取 D 而局中人 2 取 R，然后他们将进行一场猜谜博弈来确定沿路径"$D\to R\to$终点结"的盈利，因此自然地以猜谜博弈的平均盈利来代替。

猜谜博弈可以视作图 4.21 中整体博弈的一个"残剩"部分，如同二人下棋一样。从任何一着棋开始到下棋结束的过程称为"残

局”，残局本身自成博弈，我们可以称它为下棋博弈中的子博弈，猜谜博弈同样地可以看作图 4.21 的子博弈。本节介绍的子博弈完美的基本逻辑是，用 Nash 均衡的盈利替代博弈树中任意“真子博弈”（这里所说的 Nash 均衡当然属于该子博弈），然后在缩小了的博弈树上进行后退归纳以求得合理的博弈预测。

现在我们先正式地陈述子博弈概念。展开型博弈 G_T 的一个子博弈 G_s 是如下的博弈构造：

(1)G_s 拥有与 G_T 相同的局中人，尽管这些局中人中的某些人可能在 G_s 中不采取任何行动。

(2)G_s 的初始结是 G_T 的一个单结，G_s 的博弈树由这个单结，以及这个单结的所有后续结，还有在这些结之间所有的枝一起组成。

(3)在 G_s 的终点结上每一个局中人的盈利等于原博弈 G_T 在同一个终点结的盈利。

以上三个条件等于说在原博弈中的局中人集合、行动的次序以及可行行动集均在子博弈中保留。定义本身意指每一个博弈是它自身的（平凡）子博弈。博弈的非平凡子博弈称作真子博弈(proper subgame)。在完美信息博弈中，由于每一个信息集是单一的，也就是每一个决策结是单结，因此每一个决策结也一定是某个子博弈的初始结。现在我们来回忆一下完美信息博弈的后退归纳法的六条具体做法，不难发现我们得到的后退归纳解从博弈树的每个单结开始的策略剖面一定是以该单结为初始结的子博弈的 Nash 均衡，因为我们在后退到每一个决策单结时，总是为在该结有行动的局中人选取使盈利达到最高的行动。于是，非但后退归纳解是 Nash 均衡，而且它“落实”到每一个真子博弈时，其相应的策略剖面也都是 Nash 均衡。就像（下棋）竞技中一方全胜那样，完全完美信息博弈中后退归纳解在每一个局部博弈（即真子博弈）中都是“完美”的，我们称这个后退归纳解为子博弈完美均衡。将子博弈完美均衡这个概念推广到一般的展开型博弈，有如下定义：

定义 4.3 展开型博弈的一个策略剖面称为子博弈完美均衡,如果对于该展开型博弈的每一个子博弈,该策略剖面都是 Nash 均衡。

定义中的"子博弈"包括博弈本身和所有真子博弈,因此自然地子博弈完美均衡是展开型博弈本身的 Nash 均衡,子博弈完美均衡思想最能为人们所接受。打个比方,设想展开型博弈是单人作一系列决策,假如有几套方案可以完成博弈(最终效用有所不同),那么人们乐意将每一阶段都不犯错误的方案作为合理的预测。子博弈完美实质上要求策略剖面从整体来说是合理的,从每一个单结开始的局部也是合理的,那么这样的策略剖面的确是完美无缺的。

§4.5 关于后退归纳法与子博弈完美的评论

1. 后退归纳法

后退归纳的逻辑推理是令人信服的。凡能使用后退归纳的展开型博弈的预测中我们常常会想到使用它,但是,正如我们前面所指出的,后退归纳的应用具有局限性。例如,对信息的完美性与博弈的是否有限等都有一定的要求,本节将从其他的观点来评论后退归纳,先从一个小例子开始(见图 4.22)。

这是一个 n 人博弈,每个局中人 $i(<n)$在决策结上有两种选择:要么取 D 结束博弈,所有局中人各获盈利 $1/i$;要么取 A 以让下一个局中人去作选择。倘若 n 个局中人都取 A,那么每人都将满意地获益 2。图 4.22 绘制的博弈树明白地显示这是一个有限个局中人、有限策略空间、有限行动系列的完全且完美信息的展开型博弈。其后退归纳解为 $s=(A,A,\cdots,A)$,相应的盈利向量为$(2,2,\cdots,2)$。用句通俗一点的话,只要大家齐心协力一致地取行动 A,一

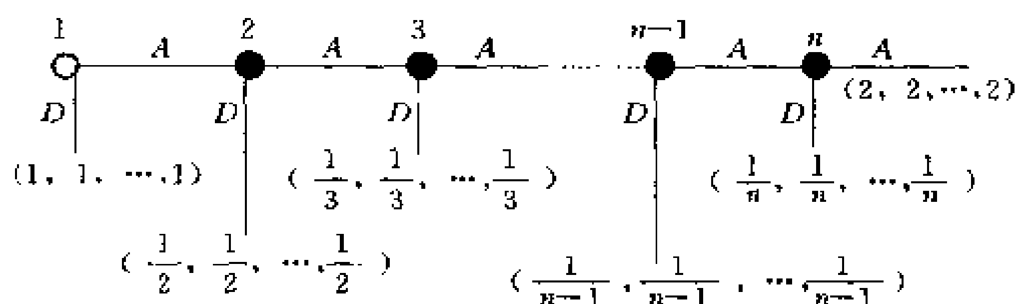


图4.22

定会有丰厚的回报。这使我们想起了完全信息静态博弈中的"共同投资"问题，它的 Nash 均衡解之一就是所有公司坚持投资大工程，以获取最大回报。除了静态与动态以及盈利函数方面的差异之外，这两类博弈很有共同相似之处。回忆共同投资问题中，存在着一个稳健性(robustness)问题，如果局中人个数 n 比较少，预测大家投资大工程的把握比较大，但若 n 比较大时，对于每一个局中人来说，只有在吃准每家公司投资大工程的可能性(概率)相当大时，才能预测"共同富裕"是合理结局。这个稳健性问题在我们的动态博弈中也具有重要意义。

不妨假设每个局中人取 A 的概率均为 p，并假设各个局中人的行动是互相独立的，至少当每个局中人在自己的决策结上选择行动时，他是在独立地思考。先观察 n 很小(例如 $n=2$)的情况，局中人 1 在 A 与 D 之间进行选择时，只需考虑到局中人 2 的可能选择，显然由后退归纳逻辑，局中人 2 取 A 或取 D 都将结束博弈，从盈利比较，他应当取 A 而不取 D，除非他不是理性的。基于这样的认识与预测，局中人 1 取 A 是明智的行动。可是，如果 n 相当大的话，局中人在行动之前非但要考虑到局中人 2 的选择，而且要考虑到另外 $(n-2)$ 个局中人的行动选择，只有从局中人 2 直到局中人 n 都取 A 时，局中人 1 取 A 才算是明智之举。根据我们的假设，发生这个前提的概率为 p^{n-1}，如果 n 比较大时，即使 p 比较大，这个概率还是比较小。譬如 $p=0.90$，$n=20$，则 $p^{19}\approx0.314$。较小的概

率很可能动摇局中人 1 取 A 的决心。同样，人们也完全可以怀疑局中人 2 会有相同的忧虑，局中人 2 有可能取 D 以预防后续的局中人会“失算”、“意外”或“犯错误”，甚至某一个局中人故意取 D（尽管博弈假设局中人都是理性的，但人一多，什么样的事情难保不会发生）等等。鉴于这些考虑，增大了局中人 1 取 D 的可能性，毕竟这样大家都可获益 1。

毛病就出在后退归纳的环节过长了一些，尽管后退归纳是从后往前合乎逻辑地推理，但是博弈却是实实在在地由局中人 1 开始逐步展开：局中人 1 取 A，需要假设局中人 2 直到局中人 n 均理性化；需要假设局中人 2 也能预测到局中人 3 至局中人 n 都会理性地取 A，又需要假设局中人 2 能够像自己预测局中人 2 知道该做什么那样地预测局中人 3 应该知道怎样做；……；等等。n 越大，这一系列的假设也越多，“夜长未免梦多”，只要其间任何一个环节出那么一点小差错或小变化，对博弈的预测结局将会非常地敏感。以生活中常发生的“传口令”游戏为例，设有 n 个士兵，在夜行军中途从头到尾每人向紧跟后面的人传达口令直到最后一个士兵听到为止。士兵们个个是理性的，传口令时人人是认真的，因此我们可以预测最后一个士兵听到的一定是走在最前面那个士兵的原话。假如 $n=3$、5 或 10，反正 n 比较小的时候，我们的预测可以保证基本正确，但若 n 过大，例如3 000甚至10 000，过长的链使我们对自己的合理预测免不了抱有担忧。

因此合理的后退归纳解可能由于整个链太长而出现其结果与博弈实际结果大相径庭的现象。正式一点地说，过长的后退归纳链可能导致博弈预测的稳健性较差。

那么，在任何可以应用后退归纳法的模型中，如果后退归纳链并不太长，后退归纳解就一定是合理的令人信服的预测结果吗？考虑如图 4.23 所示的展开型博弈。

这是一个构思十分巧妙的两人轮流行动博弈，巧妙在于各个

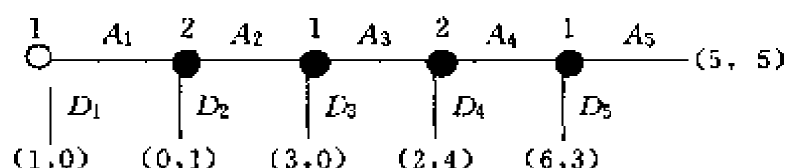


图4.23

终点结处的盈利向量，恰好使后退归纳解宣布，在每一个决策结上行动的局中人应取行动 D_i($i=1$ 或 2 或 3 或 4 或 5)。这个解令人信服吗？

作为局中人 2，如果局中人 1 在第一次行动时就发生偏离，采取了行动 A_1，他将怎么办？按照后退归纳法，他应当取 D_2，否则若取 A_2 的话将会给局中人 1 一个机会使得局中人 1 取届时的后退归纳结果 D_3。这种推论是根据"局中人 1 是理性的"这一假设作出，然而根据这一假设，后退归纳法告诉局中人 2"局中人 1 应该早就取 D_1 从而结束博弈"。摆在人们面前的问题是，一旦局中人"反常"地取 A_1 而不取 D_1，后退归纳逻辑如何帮助局中人 2 以预测局中人 1 在未来怎样行动呢？由于局中人 1 取 A_1，后退归纳逻辑说局中人 1 并非理性，既然不是理性，那么在局中人 2 若取 A_2 的情况下，局中人 1 就不能担保会按照后退归纳逻辑取 D_3(这对局中人 2 不利)，至少存在着他取 A_3 的可能(这样对局中人 2 来说，比他自己取 D_2 要好得多)。也许局中人 2 可以这样地进行逻辑推理：既然局中人 1"吃错了药"或者"犯小错误"或由于其他不详原因偏离了后退归纳解 D_1，那么在自己取 A_2 后局中人 1 仍存在一定可能不取后退归纳解 D_3，加上反正对手所取 D_1 及可能取的 D_3 对局中人 2 来说盈利均为零这一因素，因此局中人 2 倘若也违背后退归纳解 D_2 而采取行动 A_2，这不失为一次好的冒险。

经济文献中的大多数动态博弈分析继续在无保留地使用后退归纳法和它的改进，但是近几年来对后退归纳持怀疑态度者正在趋于越来越多。图 4.23 是根据 Rosenthal 于 1981 年建立的例子

进行改编而得，此君是最早质疑后退归纳逻辑的人士之一。诸如后退归纳法等博弈理论并没有为局中人在那些意外事件的一旦发生后提供预测方法与手段。Basu 等博弈论专家已经指出有关行动的合情合理的理论不应当试图排除那些在理论上为零概率的事件一旦发生后的任何行为。1988 年 Fudenberg、Kreps 与 Levine 在他们所做的工作中提出局中人把未预期到的（或意外的）偏离解释为由于实际盈利不同于原先认为最可能的盈利而引起。既然行动的任何观察结果可以用有关对手的盈利的一些说明来解释，那么这种说法回避了基于零概率事件的条件下形成局中人信念的困难，并且将"偏离"之后如何预测行动的问题彻底改变为在给定的观察到的行动下那些其他的盈利最有可能的问题。Fudenberg 与 Kreps（1988）把这种讲法推广为一种方法准则：他们提出，在指定正概率到任何可能的行动序列的意义下，任何行动理论应该是"完整的"，因此，使用该理论，局中人的关于后续行动的条件预测总是确定好的。

盈利函数的不确定性不是得到完整理论的唯一方式。第二族的完整理论通过将任何展开型博弈解释为不言明地包括了局中人有时犯小"错误"或者在 Selten（1975）意义下的"颤抖"（trembles）这样的事实。如 Selten 所假设，如果在不同的信息集上颤抖概率为独立的话，那么不管过去的行动常不能与后退归纳预测相一致，局中人在从现时开始的子博弈中继续使用后退归纳以预测行动被证明是有理的。于是用"颤抖"解释偏离是为后退归纳辩护的办法。相关的问题是局中人如何看待偏离的"颤抖"解释。具体地说，在图 4.23 中，如果局中人 2 观察到 A_1，他面临的问题是，他应当将这个观察结果解释为"颤抖"，还是将它看作局中人 1 可能取 A_3 的一个信号。

2. 子博弈完美

从子博弈完美是由后退归纳法引出的这一事实，可知子博弈完美均衡其实是后退归纳解的推广。既然后退归纳存在着一些争

议，那么殃及子博弈完美亦在情理之中。而且正由于子博弈完美的范围更宽广，因此易引起争议的内容也就多了一些。我们知道子博弈完美的基本逻辑是，以子博弈的 Nash 均衡的盈利来替代博弈树中的该子博弈，以便在缩小了的博弈树上可以后退归纳。问题就出在如果某子博弈存在多重 Nash 均衡，那么我们以哪一个 Nash 均衡来替代呢？子博弈完美均衡要求所有的局中人在处理这类问题上"心往一处想"，预测到同一 Nash 均衡，否则很有可能出现另外的"改进"均衡。图 4.24 所示的例子就是根据这种思想构造的。

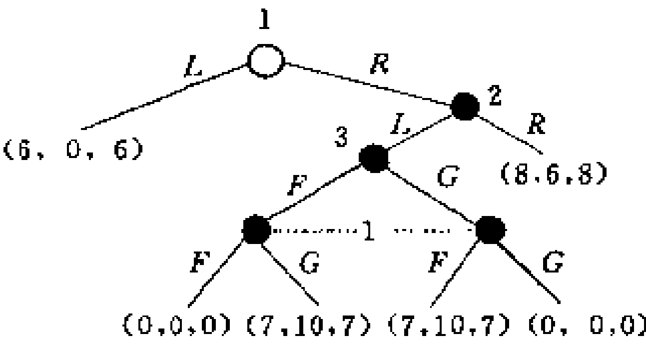


图4.24

这是一个三人展开型博弈，局中人 1 先行并有 L 与 R 两个选择，若取 L 则结束博弈，若取 R 则由局中人 2 作决策。局中人 2 也有 L 与 R 两种选择，若局中人 2 取 R 则博弈结束，但若取 L，则博弈进入第三阶段，此时由局中人 1 与 3 进行"协调博弈"，如果他们二人同时选择 F 或 G，则三个局中人都将一无所获，但如果他们二人从 F、G 中选择不同的行动，则各人得 7 而局中人 2 可获 10。

显然我们应先考察"协调博弈"，这里不必考虑局中人 2 的盈利，该博弈的策略型表示如图 4.25。

局中人 3

		F	G
局中人 1	F	0,0	7,7
	G	7,7	0,0

图 4.25

(G,F)与(F,G)是协调博弈的两个纯策略 Nash 均衡，现在令局中人 3 取 F 的概率为 p，取 G 的概率为$(1-p)$，那么局中人 1 取 F 的期望盈利为 $0\times p+7\times(1-p)$，而局中人 1 取 G 的期望盈利为 $7\times p$，令 $7\times p=7\times(1-p)$可得 $p=1/2$，由于图 4.25 中盈利矩阵关于两个局中人的对称性，易知协调博弈的混合策略 Nash 均衡为$\{(1/2,1/2),(1/2,1/2)\}$，它使得两个局中人各得期望盈利 3.5，不难算得在该均衡下，局中人 2 的期望盈利为 5。

如果局中人 1 与 3 协调成功，即取盈利向量(7,10,7) 替代子博弈，比较(7,10,7)与(8,6,8)，局中人 2 应取 L。再以(7,10,7)与(6,0,6)比较，局中人 1 应取 R。倘若以“无效”的混合策略期望盈利来替代，(3.5,5,3.5)与(8,6,8)相比，局中人 2 的决策应取 R，再以(6,0,6)与(8,6,8)相比较，显然局中人 1 应取 R。结论是，该三人博弈的所有子博弈完美均衡中，局中人 1 都应取行动 R。

Rabin 指出，局中人 1 取行动 L 仍然是合理的，假如他发现在第三阶段无法与局中人 3 成功地协调的话，他会这样去做。此时他考虑到在这个将可能到达的阶段他的期望盈利为 3.5，令他担忧的是局中人 2 将相信或预测第三阶段的协调是成功的，这样的预测使局中人 2 取 L，而若局中人 1 取 R 的话所得平均盈利 3.5 当然地小于他从一开始就取 L 时可以得到的盈利 6。显然局中人 1 取 L 不在均衡路径上，发生这样的情况，问题就在于局中人 1 与局中人 2 对协调博弈所期望的均衡不是同一的。

第五章　多阶段博弈子博弈完美的应用

§5.1　可观察行动多阶段博弈

博弈论关于经济学、政治学、管理科学与生态学的许多应用使用了特殊的一类“可观察行动的多阶段博弈”(此类博弈也常称作“几乎完美信息的博弈”)。

所谓可观察行动的多阶段博弈通常是指：

(1)在第 k 阶段中所有的局中人在选择行动时知道在此之前的 0,1,2,…,$(k-1)$阶段所选择的行动(为方便起见，我们总把初始阶段看作 0 阶段)。

(2)在每一个 k 阶段，所有的局中人“同时行动”。每一个局中人在不知道其他任何局中人在该阶段的行动时选择自己的行动。习惯上，允许如下可能性：在 k 阶段某些局中人只有单元素选择集“不做什么”。这样，“同时行动”就不排斥常见的局中人交替行动的博弈，例如二人下棋。

根据定义，由于两个企业同时选择产量，Cournot 竞争是一阶段博弈，而 Stackelberg 模型则为二阶段博弈。所谓阶段，常给人以“时间周期”或“期间”的印象。事实上，在不少场合“阶段”的确是以时间周期来确定的，然而也并不总是如此。讨价还价模型就是一个反例。在那里，每一个时间周期内存在着两个阶段，第一个阶段由局中人 1 提出价格建议，同期间的另一个阶段是局中人 2 对此建

议表示拒绝或者接受。

每一个局中人 $i(i=1,2,\cdots,n)$ 在 k 阶段选择行动前，他所了解的有关以前阶段的所有行动——即他所获得的信息集，记为 h^k。我们也称 h^k 为 k 阶段的历史，$A_i(h^k)$表示局中人 i 在 k 阶段可供选择的行动集合。如果以 α_i^j 表示局中人 i 在 j 阶段的行动，那么 n 个局中人在 j 阶段的行动可记为

$$\alpha^j=(\alpha_1^j,\alpha_2^j,\cdots,\alpha_i^j)$$

而历史 h^k 其实是 $\alpha^0,\alpha^1,\cdots,\alpha^{k-1}$的全体：

$$h^k=(\alpha^0,\alpha^1,\cdots,\alpha^{k-1})$$

毫无疑问，0 阶段的历史 h^0 应为空集，$h^0=\varnothing$，局中人 i 的一个纯策略其实对每一个 k 与每一个历史 h^k 确定一个行动 $\alpha_i\in A_i(h^k)$的映照 s_i。这与展开型博弈中纯策略的定义是一样的。

从定义来看，可观察行动的多阶段博弈符合展开型博弈的要求，它的确属于展开型博弈。可是有时出现令人“讨厌”的情况：有可能存在两个展开型式，它们似乎表示同一个博弈，但是其中一个是多阶段博弈而另一个则不是。

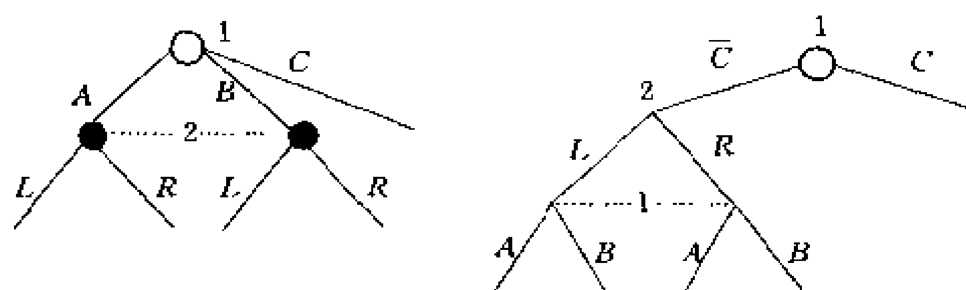


图5.1

图 5.1 左面的展开型不是一个多阶段博弈：由于局中人 2 的信息集不是单一的，依多阶段定义，它必定属于第一阶段而不是第二阶段。但是局中人 2 还拥有局中人 1 在第 1 次行动中的一些信息，譬如，如果局中人 2 的信息集到达，那么局中人 2 知道局中人 1 在第一阶段至少不取行动 c，因此局中人 2 的信息集又根本不可

能属于第一阶段。这种矛盾现象只能表明它不可能是多阶段博弈。不过，图 5.1 右边的展开型明显的是个两阶段博弈：第一阶段局中人 1 在 C 与 $\overline{C}(\overline{C}=A\cup B)$ 之间选择，一旦他选择 $\overline{C}$，则二个局中人进入同时行动的静态博弈，即第二阶段。有趣的是，这两个展开型似乎指示着同一情况：当局中人 2 行动时，他知道局中人 1 正在选择 A 或 B 而不是 C，而局中人 1 在不知道局中人 2 有关 L 或 R 的情况下选择 A 或者 B。

我们关心的是可观察行动的多阶段博弈的一个策略剖面是否子博弈完美均衡，如何去证实这件事呢？下一节介绍的一阶段偏离准则是一种办法，它是基于后退归纳基础上的关于动态规则的最优准则，有助于阐述子博弈完美如何地推广了后退归纳思想。

§5.2 有限范围博弈的一阶段偏离准则

定理 5.1 在可观察行动的有限多阶段博弈中，策略剖面 s 是子博弈完美，当且仅当 s 满足一阶段偏离条件：没有一个局中人 i 通过在某单阶段偏离然后再与 s 一致而获利。更精确地，策略剖面 s 是子博弈完美的，当且仅当不存在某局中人 i 及其策略 $\hat{s}_i$ 除了在单个 t 以及 h^t 之外，$\hat{s}_i$ 与 s_i 一致，在 h^t 到达的条件下 $\hat{s}_i$ 比 s_i 是关于 s_{-i} 的较好反应。

注：甚至可以更精确地说，不可能存在这样的历史 h^t，使得在子博弈 $G(h^t)$ 上，$\hat{s}_i$ 优于 s_i。

其实，当 s 是子博弈完美均衡的话，由定义，一阶段偏离准则中的要求自然应该得到满足。我们关心的问题主要是，如果策略剖面 s 满足定理 5.1 的条件，s 是否一定是子博弈完美均衡。也就是说，定理 5.1 的必要性是显然的，关键是要证明它的充分性。一阶段偏离准则非但有理论上的意义，在实用中也有其应用价值。例如

在阶段数较少的模型中，尤其在两阶段博弈，一阶段偏离准则常可以用来核实策略剖面是否子博弈完美。本章稍后涉及的例子中就用上了该准则。而若想论证 s 不是子博弈完美的，仅需寻求某个局中人在某阶段的小小偏离会使他获益匪浅就足够了，在操作上比较简单和方便。

采用反证法来证明定理 5.1，假定 s 满足如下条件：

(1)一阶段偏离条件：即不存在局中人 i 仅通过在某阶段偏离 s(而后又回到 s)而获得好处。

(2)但 s 不是多阶段博弈的子博弈完美均衡。

由条件(2)，必定存在一个阶段 t 及历史 h^t，使得某局中人 i 有一个策略 $\hat{s}_i$，在从信息集 h^t 开始的子博弈中，$(\hat{s}_i, s_{-i})$ 优于 (s_i, s_{-i})。显然可以设想 $\hat{s}_i$ 与 s_i 在 h^t 处不相同，但是由一阶段偏离条件(1)，$\hat{s}_i$ 与 s_i 不可能只在 h^t 不同，不妨先假设在 h^t 开始的子博弈中还恰好存在另一个历史 $h^{\hat{t}}$，使得 $\hat{s}_i(h^{\hat{t}})\neq s_i(h^{\hat{t}})$，显然 $\hat{t}>t$。由于博弈的有限性，取 $\hat{t}$ 为有限。

现在我们建立一个新的 $\bar{s}_i$ 如下

$$\bar{s}_i=\begin{cases}\hat{s}_i & \text{当 } t^*<\hat{t}\\ s_i & \text{当 } t^*\geqslant\hat{t}\end{cases}\qquad(5.1)$$

注意到 $\hat{s}_i$、s_i、$\bar{s}_i$ 从 $\hat{t}+1$ 开始全都一样。在从 $\hat{t}$(即 $h^{\hat{t}}$)开始的每一个子博弈上，由于仅有 $\hat{s}_i(h^{\hat{t}})\neq s_i(h^{\hat{t}})$而从 $\hat{t}+1$ 开始 $\hat{s}_i$ 与 s_i 相同，由一阶段偏离条件，$\hat{s}_i$ 是与 s_i 一样好的反应，故 $\bar{s}_i$ 与 $\hat{s}_i$ 也是同样好的反应(因为 $\bar{s}_i=s_i$，当 $t^*\geqslant\hat{t}$)。而在 $t^*<\hat{t}$ 时有 $\bar{s}_i=\hat{s}_i$，因此在从 t 开始具有历史 h^t 的子博弈中，$\bar{s}_i$ 与 $\hat{s}_i$ 也是同样好的反应。由于 $\bar{s}_i$ 的定义，它与 s_i 仅在 h^t 时不同，再由一阶段偏离条件知，在从 t 开始且具有历史 h^t 的子博弈中，$\bar{s}_i$ 应该与 s_i 一样好的反应。从而 $\hat{s}_i$ 也是与 s_i 一样好的反应。这与 $\hat{s}_i$ 是优于 s_i 的反应矛盾。所以由一阶段偏离条件可以推断 s 还满足下述条件。

(3)一、二阶段偏离条件：即不存在局中人 i 通过在某一个阶

段或在某二个阶段偏离 s，在其他阶段仍回到 s，而从中获益。

得到(3)并不等于已经证明了定理，如果我们由一阶段偏离条件推得，不存在局中人 i 通过在任意有限个阶段偏离 s（其余阶段回到 s）而从中获益。这才算真正证明了定理。现在我们不妨假设 $\hat{s}_i$ 除了在 h^t 处与 s_i 不同，在 t 以后还有两个阶段也存在不等式 $\hat{s}_i(h^{t'})\neq s_i(h^{\hat{t}})$。以 $\hat{t}$ 记这两个 t' 中的最大者。再构造局中人 i 的新策略 $\bar{s}_i$ 如下：

$$\bar{s}_i=\begin{cases}\hat{s}_i & \text{当 } t^* < \hat{t}-1 \\ s_i & \text{当 } t^* \geqslant \hat{t}-1\end{cases} \tag{5.2}$$

从$(\hat{t}-1)$开始的子博弈上，$\bar{s}_i$ 与 s_i 至多在 $h^{\hat{t}-1}$上不同，由(1)，$\bar{s}_i$ 与 s_i 一样地好。而 $\hat{s}_i$ 与 s_i 至多在 $h^{\hat{t}-1}$与 $h^{\hat{t}}$上不同，由(3)，$\hat{s}_i$ 与 s_i 一样好。因此易知 $\bar{s}_i$ 是与 $\hat{s}_i$ 一样好的反应。由于当 $t^* < (\hat{t}-1)$时 $\bar{s}_i=\hat{s}_i$，因此在从 t 开始具有历史 h^t 的子博弈上 $\bar{s}_i$ 与 $\hat{s}_i$ 一样好。再由(5.2)式，$\bar{s}_i$ 与 s_i 在从 t 开始的子博弈中至多在 2 处历史上有所不同。利用(3)，$\bar{s}_i$ 与 s_i 一样地好，因而 $\hat{s}_i$ 与 s_i 一样地好，这与 $\hat{s}_i$ 的假设矛盾。这样我们从一阶段偏离条件又可进一步推断到“一、二、三阶段偏离条件”，从该条件出发，再考虑 $\hat{s}_i$ 与 s_i 有 4 处不同的情况，利用上述方式进行推理，一步一步地“推广”一阶段偏离条件，由于研究的是有限多阶段博弈，因此利用数学归纳原理，经过有限步“推广”之后即可证得定理 5.1。

上述证明并没有排斥如下可能性，尽管局中人 i 不能通过一阶段偏离而在任意子博弈中获益，但是他可能通过无限序列的偏离而获益。这个问题涉及到无限博弈的一阶段偏离准则，本书将略作介绍。

定理 5.2 （无限多阶段博弈的一阶段偏离准则）在可观察行动的无限多阶段博弈中，如果博弈“在无穷远处连续”，那么策略剖面 s 是子博弈完美均衡，当且仅当，不存在局中人 i 和策略 $\hat{s}_i$（$\hat{s}_i$ 与

s_i 仅在单个 t 的历史 h^t 上不同，其余的历史上 $\hat{s}_i$ 与 s_i 完全一致），使得在历史 h^t 到达的条件下，$\hat{s}_i$ 比起 s_i 来，是关于 s_{-i} 的较好反应。

定理 5.2 比定理 5.1 多了一个"在无穷远处连续"的条件，必须对此有所交代。在无限多阶段博弈中，常令 h 表示无限博弈的一个结局，其实，该 h 也代表某个无限水平的历史。对每一个历史 h，令 h^t 表示 h 关于首 t 个阶段的部分，也就是相应于 h 在 t 阶段时的历史。对于局中人 i，关于历史 h（或结局 h）的盈利可记为 $u_i(h)$。对于不同的无限水平历史 h_1 与 h_2、相应地有不同的盈利 $u_i(h_1)$ 与 $u_i(h_2)$，如果 h_1 与 h_2 在 t 阶段具有相同的历史 $h_1^t=h_2^t$，由于 h_1 未必一定就是 h_2，$u_i(h_1)$ 与 $u_i(h_2)$ 之间所存在的绝对差异可表示为：

$$(|u_i(h_1)-u_i(h_2)| \parallel h_1^t=h_2^t) \doteq \Delta(i,h_1,h_2,t) \qquad (5.3)$$

使得 $h_1^t=h_2^t$ 的历史"对"(h_1,h_2) 可以有许多，对这些 $\Delta(i,h_1,h_2,t)$ 求最大者（如果"对"数无限多，则取上确界）：

$$\sup_{h_1,h_2}\Delta(i,h_1,h_2,t) \qquad (5.4)$$

如果对每一个局中人 i，当 $t\to\infty$（因为是无限多阶段博弈，故 t 可趋向无穷），我们有

$$\lim_{t\to\infty}\sup_{h_1,h_2}\Delta(i,h_1,h_2,t)=0 \qquad (5.5)$$

那么该无限多阶段博弈称为在无穷处连续。从前面的详细分析及(5.5)可知，如果 t 相当大，对每一个局中人 i 来说，不论无限水平历史 h_1 与 h_2 怎么样，只要 $h_1^t=h_2^t$，那么他在这两个结局 h_1、h_2 的盈利几乎相差很小。换句话说，"在无穷处连续"意指对于局中人的盈利，很远的将来事件相对地是不重要的。

条件"在无穷处连续"在如下情况常能满足：如果每个局中人 i 在无限多阶段博弈中的总体盈利可以表示为每阶段（或周期）盈利的折扣和，且每周期的盈利为一致有界。这种情况在讨论重复博弈时常会遇到。

现在我们来证明无限多阶段博弈的一阶段偏离准则。定理的必要性仍然从子博弈完美均衡的定义立即可得，而且在证明定理5.1的充分性时我们实际上证明了，如果 s 满足一阶段偏离条件，那么不可能通过任意子博弈中任何有限次偏离而对 s 进行改善，此结果我们记为(＊)。

现在仍用反证法，设 s 满足一阶段偏离条件，但 s 不是子博弈完美均衡。那么将存在一个阶段 t 和在 t 处的历史 h^t，在那里某局中人 i 通过在 h^t 开始的子博弈中使用一个不同的 $\hat{s}_i$ 策略来增加他的盈利。不妨设增加的盈利量为 $\varepsilon>0$，即 $\hat{s}_i$ 与 s_i 从 t 开始不同，而在 t 之前相同，且有

$$u_i(\hat{s}_i)=u_i(s_i)+\varepsilon \tag{5.6}$$

由于博弈是在无穷远处连续，对于给定的 $\varepsilon>0$，总存在充分大的 $t'(>t)$，使得在 t' 之前与 $\hat{s}_i$ 一样的策略 $\bar{s}_i$ 与 $\hat{s}_i$，它们各自所导致的盈利之间的差异小于 $\varepsilon/2$。该策略 $\bar{s}_i$ 可以定义如下：

$$\bar{s}_i=\begin{cases}\hat{s}_i & t^*<t' \\ s_i & t^*\geqslant t'\end{cases} \tag{5.7}$$

且 $\bar{s}_i$ 必满足

$$u_i(\hat{s}_i)-\frac{\varepsilon}{2}\leqslant u_i(\bar{s}_i)\leqslant u_i(\hat{s}_i)+\frac{\varepsilon}{2} \tag{5.8}$$

因此在 h^t 开始的子博弈中，$\bar{s}_i$ 仅在 t 至 $(t'-1)$ 之间有限个阶段可能与 s_i 不同，但总有

$$u_i(\bar{s}_i)>u_i(\hat{s}_i)-\frac{\varepsilon}{2}=u_i(s_i)+\frac{\varepsilon}{2} \tag{5.9}$$

(5.9)式表明我们通过在 h^t 开始的子博弈中作有限阶段偏离而对 s 进行了改善，这与前面所述结论(＊)矛盾。从而证明了定理5.2中一阶段偏离条件的充分性。

§5.3 具有静态均衡的重复博弈

重复博弈在经济上有许多应用,它在博奕论中占有重要地位,我们将另立一章专门进行讨论。这里我们初步涉及重复博弈完全是因为,一则它是一类特殊的可观察行动的多阶段博弈,二则我们可以在重复博弈上试试刚介绍的"一阶段偏离准则"的功效究竟如何。

从我们所熟悉的囚徒窘境谈起,不过为了方便起见,将效用矩阵改写如图 5.2 所示。

		囚徒2 合作	囚徒2 背叛
局中人1	合作	1,1	−1,2
	背叛	2,−1	0,0

图 5.2

这里的"合作"、"背叛"是指两个囚徒互相之间的关系,原效用矩阵中的元素是指关押时间,现在的盈利矩阵可以解释为他们的收入。例如有两个小偷,如果他们互相合作拒不交代的话,各人保持偷得的一份财物;如果互相揭发坦白,那么偷得东西被没收,盈利为 0;如果一方坦白一方抗拒,那么坦白的一方非但不要交出财物反而得到一份奖励,抗拒的一方不仅充公偷物还要罚出一份东西,当然这一切是虚拟假设的。假定囚徒窘境每年一次地重复 $(k+1)$次,显然这是一个可观察行动的$(k+1)$阶段博弈。设 t 阶段的行动向量为 a^t,局中人 i 在该阶段博弈盈利为 a^t的函数 $g_i(a^t)$。$(k+1)$次重复博弈的总盈利相当于每年盈利的总和。这里涉及到一个折扣因子问题。设想今年的 1 元钱如果存入银行,令银行的一年利息为 γ,那么到明年就变成$(1+\gamma)$元。从今年的观点看,明年

的 1 元盈利相当于今年的 $\delta=1/(1+\gamma)$ 元，显然，$0<\delta\leqslant 1$，这个 δ 就是明年的折扣因子，当然每年的折扣因子可以不同。现在假定每年的折扣因子都是 δ，从第 1 年的观点，共 $(k+1)$ 次博弈的总盈利（对局中人 i）应为

$$g_i(\alpha^0)+\delta g_i(\alpha^1)+\delta^2 g_i(\alpha^2)+\cdots+\delta^k g_i(\alpha^k)=\sum_{t=0}^{k}\delta^t g_i(\alpha^t) \tag{5.10}$$

这相当于 $(k+1)$ 年总盈利的“现时价值”。倘若每次盈利为 1，那么总盈利的现时价值为 $(1-\delta^{k+1})/(1-\delta)$ 而不是 $(k+1)$。将这件事说成每年平均盈利为 1 是合理的，基于这种想法，于是我们定义“平均折扣盈利”为

$$\frac{1-\delta}{1-\delta^{k+1}}\sum_{t=0}^{k}\delta^t g_i(\alpha^t) \tag{5.11}$$

这种做法使得不同水平（即不同的 k）的盈利可以互相比较。

若取 $k=0$，即只实施一次囚徒窘境，显然“合作”为严劣策略，（背叛，背叛）是该博弈的唯一 Nash 均衡。如果该博弈再重复一次，那么可以运用后退归纳的原理，先看后一次囚徒窘境，它的 Nash 均衡是（背叛，背叛），后退到第一次囚徒窘境，其 Nash 均衡仍为（背叛，背叛），由一阶段偏离准则，不难知道{（背叛，背叛），（背叛，背叛）}是重复囚徒窘境的子博弈完美。同样的推理，囚徒窘境只要重复有限次，那么每一次双方互相背叛构成了唯一的子博弈完美均衡。

如果囚徒窘境（图 5.1）无限次地进行，后退归纳的思想无法照此实行。但是，无限水平博弈的一阶段偏离准则将告诉我们“每次双方均背叛”仍然是子博弈完美均衡。因为每一次互相背叛使各人盈利为零。由于两囚徒是在隔离情况下独立地作出决策的，在任何一次博弈中谁也不敢发生偏离而取“合作”（即抗拒交代），因为这会给他带来（-1）盈利。人们常常会想到，在无限水平的博弈中，

前一阶段的结果将会影响到后面阶段的决策，而这里的子博弈完美均衡却是唯一地具有这样的性质：每一阶段的行动并不随以前阶段所采取的行动而变化。但是，其他的子博弈完美均衡却不具有这种性质，例如当折扣因子 $\delta>\frac{1}{2}$ 时，我们有如下的子博弈完美均衡："在第一个周期内双方合作，只要没有一个局中人在任何时候背叛，那么就继续合作下去。如果任何一个局中人在某阶段背叛，那么博弈的以后部分双方均互相背叛。"显然，这个策略剖面存在着前面阶段的行动影响后面阶段决策的现象。现在我们试图证明它的确是子博弈完美。就是要检验一下是否存在一个子博弈，某局中人只偏离一下会给他带来更好的盈利。就该策略剖面而言，存在着两类子博弈：

A 类——没有发生局中人背叛的事件；

B 类——发生了背叛并一直背叛下去。

如果一个局中人与 A 类中每一个子博弈的策略一致，那么他的平均折扣盈利显然为 1，这是因为

$$1+\delta+\delta^2+\cdots=1/(1-\delta)$$

因而

$$\frac{1-\delta}{1}\sum_{t=0}^{\infty}\delta^t=\frac{1-\delta}{1-\delta}=1$$

在无限水平博弈中，(5.11)式中求和号前的系数应为 $1-\delta$。假如该局中人在 t 阶段偏离一下再回到原来的策略剖面，则他从 t 阶段开始转人了 B 类子博弈。那么他的平均折扣盈利不再是 1 而是

$$(1-\delta)(1+\delta+\cdots+\delta^{t-1}+2\delta^t+0+0+\cdots)=1-\delta^t(2\delta-1) \tag{5.12}$$

当 $\delta>\frac{1}{2}$ 时，(5.12)式显然小于 1。这表明局中人若处于 A 类，他不可能通过一阶段偏离获得更多盈利。现考虑 B 类情况，对于 B 类中的任意历史 h^t，与从 t 周期开始的策略一致的盈利都是 0，如

果局中人在其间偏离一次后，他又回到背叛行动并一直下去，那么他在偏离周期盈利为(−1)而在其余周期仍为 0，这对他没有好处。因此对于那个已经确定了的策略剖面，没有一个局中人能通过一阶段偏离而从中获益，根据一阶段偏离准则，这些策略形成了一个子博弈完美均衡。

上面 $\delta>\frac{1}{2}$ 并非规定，而是根据阶段博弈(即重复博弈中的每一次博弈)的盈利矩阵而确定的。如果读者愿意适当改变某些盈利大小(在不影响静态博弈的基本内容前提下)就可以得到不同的 δ 范围。

有限重复囚徒窘境的多阶段博弈，以每阶段均取 Nash 均衡形成其子博弈完美，那是因为该阶段博弈只有唯一的 Nash 均衡，利用后退归纳的思路不难断定子博弈完美的形式。倘若阶段静态博弈存在不止一个 Nash 均衡，情况很可能发生变化。例如图 5.3 所示的阶段静态博弈若实施两次，其结果就会出现微妙变化。

		局中人 2		
		L	M	R
	U	0,0	3,4	6,0
局中人 1	M	4,3	0,0	0,0
	D	0,6	0,0	5,5

图 5.3

划线法告诉我们(M,L)与(U,M)是纯策略 Nash 均衡，其相应的盈利向量分别为(4,3)与(3,4)。是否存在混合策略 Nash 均衡呢？从局中人 1 的角度出发，策略 D 的盈利向量(0,0,5)弱劣于策略 U 的盈利向量(0,3,6)，他在考虑混合策略时，置正概率于 D 不如将此正概率赋予 U 策略；同样由于盈利矩阵的对称性，局中人 2 也不会将正概率赋予纯策略 R。于是对混合策略均衡的求解，仅需考虑如图 5.4 所示的“简化”了的盈利矩阵。

		局中人 2	
		L	M
局中人 1	U	0,0	3,4
	M	4,3	0,0

图 5.4

对于图 5.4 博弈的混合策略均衡解，我们可以熟练地得到为 $\{(\frac{3}{7}U,\frac{4}{7}M),(\frac{3}{7}L,\frac{4}{7}M)\}$，相应的盈利为$(\frac{12}{7},\frac{12}{7})$。有趣的是，有效盈利(5,5)不能由均衡达到。在二阶段博弈中，假定局中人的总盈利为平均折扣盈利，而折扣因子 $\delta>\frac{7}{9}$的话，那么下述策略剖面是子博弈完美均衡："在第一阶段取(D,R)。如果第一阶段的结局是(D,R)，则在第二阶段取(M,L)；如果第一阶段的结局不是(D,R)，则在第二阶段取$\{(\frac{3}{7}U,\frac{4}{7}M),(\frac{3}{7}L,\frac{4}{7}M)\}$"。

根据该策略剖面的构造，这些策略在第二阶段确定的(M,L)或$\{(\frac{3}{7}U,\frac{4}{7}M),(\frac{3}{7}L,\frac{4}{7}M)\}$均为 Nash 均衡，因此仅需考虑在第一阶段发生偏离的可能情况。例如对局中人 1，如果他在第一阶段偏离(D,R)的话，他必定舍 D 而取 U，使自己的盈利由 5 增加到 6，增加量为 1。但是根据策略剖面构造，这一偏离被观察到之后，在第二阶段应取混合策略，局中人 1 与 2 在该阶段的盈利分别由 4 与 3 下降到$\frac{12}{7}$，也即分别下降了 $4-\frac{12}{7}$与 $3-\frac{12}{7}$，考虑到折扣因子，减少量的现时价值为 $\delta(4-\frac{12}{7})$与 $\delta(3-\frac{12}{7})$，对局中人 1，只有 $1>\delta(4-\frac{12}{7})$他才会偏离，或者说，当 $\delta>\frac{7}{16}$时局中人 1 不会偏离。而对于局中人 2，如果他在第一阶段偏离也将增加盈利 1，因此只有 $1>\delta(3-\frac{12}{7})$他才会偏离，或者说，当 $\delta>\frac{7}{9}$时局中人 2 才有积

极性在第一阶段偏离。因此，当 $\delta > \max(\frac{7}{9}, \frac{7}{16}) = \frac{7}{9}$ 时，策略剖面是 Nash 均衡，从而也是子博弈完美均衡。

§5.4 两阶段博弈的若干经济应用

1. 银行挤兑

两个投资者各具银行存款 D，银行将这两笔存款用于一长期项目。如果在项目到期之前银行被迫抽回资金，仅可挽回 $2r$，其中 $D > r > \frac{D}{2}$。倘银行同意到期后再收回，连本带息将得到 $2R$，当然 $R > D$。

客户从银行提取存款有两个日期，日期 1 为银行投资到期之前；日期 2 则为到期之后。为讨论方便起见，不考虑贴现。若两客户在日期 1 提取存款则每人各得 r 并结束博弈；如果仅有一人在日期 1 收回资金，则该客户得全款 D，余下 $2r-D$ 则为另一客户所得，此时也结束博弈；最后，若没有一个客户在日期 1 提取存款，等到项目完工后的日期 2，如果两客户均抽回资金，则每人得 R 并结束博弈；倘若只有一个客户提取资金，他可得 $2R-D$，另一客户则得 D，结束博弈；如果没有一个客户在日期 2 提取存款，银行返回每个客户 R 并最终结束博弈。

这里可以将模型视作“日期 1”与“日期 2”两个阶段，由问题表达的意思，每一阶段两客户同时行动，各人在各阶段的行动空间均为{提取，不提}，但两个阶段的盈利不一样，故不是重复博弈。注意，整个过程中，银行不是博弈的局中人，它的所有行为全随客户的行动而确定。根据这样的分析，我们不难用展开型来表示博弈(见图 5.5)。

图 5.5 中，枝旁所标“Y”与“N”分别表示行动“提取”与“不提”。利用后退归纳的思想，先考虑日期 2 时的完全信息静态博弈。

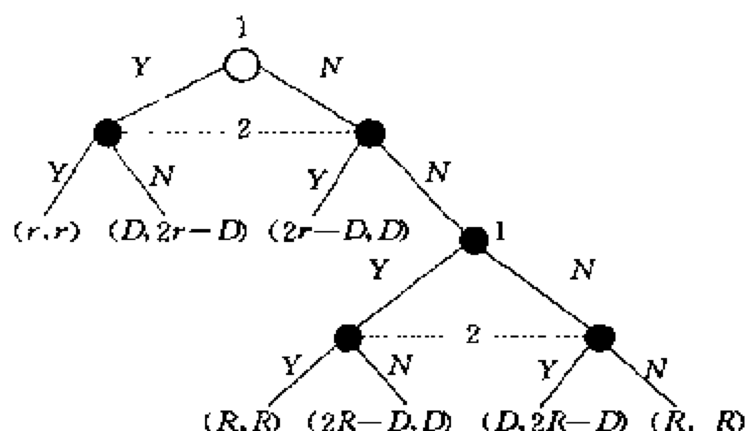


图5.5

由于 $R>D$，易知 $2R-D>R$。因此无论是客户 1 还是客户 2，"提取"严格地优于"不提"，于是得到该博弈的唯一 Nash 均衡：(提取，提取)，其盈利为 (R,R)。以 (R,R) 作为日期 2 阶段博弈的结局，得博弈树如图 5.6 所示。

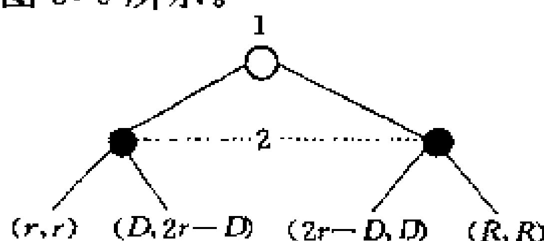


图5.6

因为 $r<D$，因此 $2r-D<r$。图 5.6 的策略型表示如图 5.7 所示。

		局中人 2	
		提取	不提
局中人 1	提取	r,r	$D,2r-D$
	不提	$2r-D,D$	R,R

图 5.7

简单的划线法告诉我们，该博弈有两个纯策略 Nash 均衡：(1)(提取，提取)导致盈利 (r,r)；(2)(不提，不提)导致盈利 (R,R)。这样，两阶段银行挤兑博弈至少有两个子博弈完美均衡：

(1)两个客户在日期 1 都提取存款，各获 r；

(2)两个客户在日期1都不提取存款，而到了日期2又都要求收回资金，此时各获得 R。

第一种结局其实就是往银行挤兑。由于种种原因，如果客户1相信客户2将在日期1提取存款，那么他的最佳反应一定是也去提取存款，即使他们都知道等到日期2再去提取，情况会好得多。银行挤兑博弈似乎与囚徒窘境一样，它们都有着导致社会无效获益的 Nash 均衡；但是它们在一个重要的方面有所区别，那就是囚徒窘境的这个无效均衡是唯一的，而银行挤兑模型却还存在着第二个 Nash 均衡，而且它是有效的。造成的后果是，这个模型不能预测什么时候将发生银行挤兑，但是揭示了作为一种均衡现象，银行挤兑有可能发生。

2. 关税与不完全国际竞争

考虑两个(虚拟的)完全相同的国家，不妨记作 $i=1,2$。每个国家都有一个政府与一个企业，由政府选择关税率，企业的产品既内销又向对方国家出口，各国消费者从国内市场购买国产货或来自对方国家的进口货。记 Q_i 为在 i($i=1$ 或 2)国市场上该产品的总量，设市场出清价格是 $P_i(Q_i)=a-Q_i$。i 国的企业(今后称企业 i)为国内消费生产 h_i 产品，并生产 e_i 供出口。显然 $Q_i=h_i+e_j(i\neq j)$。两企业均具常数边际成本 c，但无固定成本。于是企业 i 产品的总成本为 $C_i(h_i,e_i)=c(h_i+e_i)$。企业还负有出口方面的关税成本，如果企业 i 向 j 国出口 e_i，j 政府设置关税率 t_j，那么企业 i 必须付给 j 政府 t_je_i。

博弈的时序如下：首先，两个政府同时选择关税率 t_1 与 t_2；然后，企业在观察到关税率后同时选择各自的内销量与出口量(h_1,e_1)与(h_2,e_2)。显然这是一个可观察行动两阶段博弈，第一阶段有两个局中人，而第二阶段有另外两个局中人。它们的盈利或效用，对企业来说为利润，对政府而言是福利。i 国的总福利应为 i 国消费者享有的消费者盈余、企业 i 所得利润以及由 i 政府向企业 j 征

收得税款之和。

$$w_i(t_i,t_j,h_i,e_i,h_j,e_j)=\frac{1}{2}Q_i^2+\pi_i(t_i,t_j,h_i,e_i,h_j,e_j)+t_ie_j \tag{5.13}$$

其中，$\pi_i(t_i,t_j,h_i,e_i,h_j,e_j)$为企业 i 的所得利润。

$$\begin{aligned}&\pi_i(t_i,t_j,h_i,e_i,h_j,e_j)\\&=[\alpha-(h_i+e_j)]h_i+[\alpha-(e_i+h_j)]e_i-c(h_i+e_i)t_je_i\end{aligned} \tag{5.14}$$

盈利函数(5.13)式、(5.14)式具有连续统，因此博弈有点类似 Stackelberg 模型，只不过在那里局中人只有 2 人。后退归纳思想让我们先考虑第二阶段的 Nash 均衡，也就是假定两个政府已经选定了 t_1 与 t_2，$(h_1^*,e_1^*,h_2^*,e_2^*)$是我们求得的解，那么对每一个 i，(h_i^*,e_i^*)必须从下式解得：

$$\max_{h_i,e_i\geqslant 0}\pi_i(t_i,t_j,h_i,e_i,h_j^*,e_j^*) \tag{5.15}$$

注意到(5.15)式在给定 h_j^* 与 e_j^* 下，可以写成

$$\begin{aligned}&\pi_i(t_i,t_j,h_i,e_i,h_j^*,e_j^*)\\&=h_i[\alpha-(h_i+e_j^*)-c]+e_i[\alpha-(e_i+h_j^*)-c-t_j]\end{aligned} \tag{5.16}$$

第一部分表示企业 i 在国内市场的利润，第二部分则是它从出口 j 国获得的利润。这两个式子一个仅包含 h_i 而另一个仅含 e_i，因此极大化解只需对这两个市场分别求解 h_i^*,e_i^*：

$$\max_{h_i\geqslant 0}h_i[\alpha-(h_i+e_j^*)-c] \tag{5.17}$$

$$\max_{e_i\geqslant 0}e_i[\alpha-(e_i+h_j^*)-c-t_j] \tag{5.18}$$

于是，当 $e_j^*\leqslant\alpha-c$ 时，有 $h_i^*=\frac{1}{2}(\alpha-e_j^*-c)$；而当 $h_j^*\leqslant\alpha-c-t_j$ 时，有 $e_i^*=\frac{1}{2}(\alpha-h_j^*-c-t_j)$。对应地，当 $e_i^*\leqslant\alpha-c$ 时，有 $h_j^*=\frac{1}{2}(\alpha-e_i^*-c)$；而当 $h_i^*\leqslant\alpha-c-t_i$ 时，有 $e_j^*=\frac{1}{2}(\alpha-h_i^*-c-t_i)$。四

个方程有四个未知量，联立解之得

$$h_i^* = \frac{\alpha - c + t_i}{3}, e_i^* = \frac{\alpha - c - 2t_i}{3} \quad (i = 1, 2) \tag{5.19}$$

注意一个事实，(5.19)式的结果是在诸如 $e_j^* \leqslant \alpha - c$ 等假设条件下才求得的，必须关心的是所解得的结果是否与这些假设条件存在矛盾，检查(5.19)式易知

$$e_i^* = \frac{\alpha - c - 2t_j}{3} < \frac{\alpha - c}{3} < \alpha - c \tag{5.20}$$

同样有

$$e_j^* = \frac{\alpha - c - 2t_i}{3} < \frac{\alpha - c}{3} < \alpha - c \tag{5.21}$$

由(5.18)式，$e_i^* \geqslant 0$，$e_j^* \geqslant 0$，结合(5.20)式、(5.21)式有

$$\alpha - c \geqslant 2t_j \quad 或 \quad 2t_i \tag{5.22}$$

(5.22)式有其实际意义，如果政府将关税率定得过高，那么两国之间的交易将无法进行。由(5.22)式，考察 h_i^*：

$$h_i^* = \frac{\alpha - c + t_i}{3} = \alpha - c - t_i - \frac{2}{3}(\alpha - c - 2t_i) \leqslant \alpha - c - t_i \tag{5.23}$$

同理 $h_j^* \leqslant \alpha - c - t_j$

可见，$(h_1^*, h_2^*, e_1^*, e_2^*)$与为求这些解所假设的条件一致。

对第二阶段博弈的分析与结果略加讨论，如果特令 $t_1 = t_2 = 0$，即两国不设关税，那么问题立即转为国内与国外两个市场的 Cournot 竞争，所得结果与 Cournot 竞争中的均衡产量相符合，Cournot 结果是在对称边际成本假设下得到，而我们这里的博弈，由于政府的关税选择使得边际成本为非对称。例如，在 i 国市场，企业 i 的边际成本是 c，企业 j 的边际成本却为 $c + t_i$，企业 j 的成本高于企业 i，因而它打算少生产出口 i 国的产品量。然而，如果企业 j 计划少生产出口量，必将提高 i 国的市场出清价，从而企业 i 希望多生产一些，反过来企业 j 则打算生产更少出口量。于是，在均衡状态，h_i^* 随 t_i 而增加，e_j^* 则以更快速度随 t_i 而减少。可见，关

税具有保护本国企业的重要作用。

现在我们后退到第一阶段两国政府之间同时选择关税率的完全信息博弈，它们的盈利函数为 $W_i(t_i,t_j,h_1^*,e_1^*,h_2^*,e_2^*)(i=1,2)$，而 h_i^*,e_i^* 则为 t_1、t_2 的函数，从而盈利函数干脆可以简记为 $W_1^*(t_1,t_2)$ 与 $W_2^*(t_1,t_2)$。寻求该博弈的 Nash 均衡解 (t_1^*,t_2^*)，实际上是在给定 t_j^* 条件下求使 $W_i^*(t_i,t_j)$ 极大化的 t_i^*（当然求得的 t_i^* 只有在 $t_i^*\geqslant 0$ 情况下有实际意义），我们有

$$W_i^*(t_i,t_j^*)=\frac{[2(\alpha-c)-t_i]^2}{18}+\frac{(\alpha-c+t_i)^2}{9}+\frac{(\alpha-c-2t_j^*)^2}{9}+\frac{t_i(\alpha-c-2t_i)}{3} \tag{5.24}$$

易得

$$t_i^*=\frac{\alpha-c}{3} \qquad (i=1,2) \tag{5.25}$$

有趣的是，在求 t_i^* 以使(5.24)式极大化的一阶条件下，t_i^* 不依赖于 t_j^*，同理，t_j^* 也不依赖于 t_i^*。在我们建立的政府关税博弈中，在边际成本给定情况下，两国政府独立地选取自己的最优策略，而不受到他国政府选择的影响。由(5.25)式我们得到在第二阶段中的均衡量应为

$$h_i^*=\frac{4(\alpha-c)}{9},\quad e_i^*=\frac{(\alpha-c)}{9} \qquad (i=1,2)$$

于是后退归纳解$\left\{t_1=t_2=\frac{(\alpha-c)}{3},h_1=h_2=\frac{4(\alpha-c)}{9},e_1=e_2=\frac{\alpha-c}{9}\right\}$是关税问题的子博弈完美均衡。

子博弈完美均衡结局中两国市场各自的合计量均为 $5(a-c)/9$，而若政府均选择零关税率时，各个市场的合计量为 $2(a-c)/3$，恰如 Cournot 均衡一样。由于关于某市场的消费者盈余等于该市场的合计量的平方的一半，显然从社会效应来说，政府选择零关税率将优于政府选择它的占优策略关税。在这个意义下，$t_1=t_2=0$

其实是

$$\max_{t_1,t_2\geqslant 0}\{W_1^*(t_1,t_2)+W_2^*(t_1,t_2)\} \tag{5.26}$$

的解。因此对政府来说，存在一种激励以让它签署一项条约，允许零关税(即允许自由贸易)。如果允许负关税——即政府给予补贴——的话，政府的社会效应最佳选择是 $t_1=t_2=-(a-c)$，此时会引起国内企业不生产内销量，而出口到其他国家用于竞争。于是，在给定了第二阶段企业 1 与 2 采取 Nash 均衡策略的条件下，在政府之间的第一阶段的相互作用是一个囚徒窘境：唯一的 Nash 均衡是在社会效应上为无效的。

3. 竞选

考虑两个工人与他们的老板。工人 $i(i=1,2)$ 的产量为 $y_i=e_i+\varepsilon_i$，e_i 为该工人所作出的努力，ε_i 则为随机干扰，即存在着随机因素影响他的产量。生产如下进行：首先，两个工人同时选择各自的努力水平：$e_i\geqslant 0(i=1,2)$；其次，扰动 ε_1 与 ε_2 独立地来自均值为零的分布密度；第三，可观察到的是工人的产量而不是他们的努力水平。因此，工人的工资多少依赖于他们的产量而不直接依赖于他们的努力程度。

Lazear 与 Rosen 于 1981 年研究过这样的模型：老板决定通过使工人竞赛，从而激励工人努力工作。高产量的工人获得工资 w_H，产量低的工人挣得工资 $w_L(w_H>w_L)$。工人从挣得的工资以及为此努力 e 而作出的花费 $g(e)$ 中得到的综合盈利为 $u(w,e)=w-g(e)$，$g(e)$ 显然应该是 e 的递增函数，越是努力，花在该努力水平上的费用也就越高。对于不同的努力水平 e_1 和 e_2，相应的花费是 $g(e_1)$ 与 $g(e_2)$，依据常理，$(e_1+e_2)/2$ 的花费通常应当不大于 $[g(e_1)+g(e_2)]/2$，因此我们假定 $g(e)$ 是凸函数。老板的盈利比较简单，假如工人 1 与 2 分别为他生产的价值为 y_1 与 y_2，那么这两个数之和再减去必须支付给工人的工资就是老板的盈利：y_1+y_2

$-w_H-w_L$。

现在我们将模型用两阶段博弈来叙述。第一阶段可以视作老板的“单人博弈”，即作出决策——选择竞赛中要付出的工资 w_H 与 w_L。在第二阶段，两个工人在观察到 w_H 与 w_L 之后进行博弈，他们同时选择努力水平 e_1 与 e_2（以后我们还会考虑这样的可能性，根据老板确定的工资，工人不愿参加而宁可接受另外的职业）。由于产出（同时也是工资）不仅是局中人行动的函数而且还是随机干扰 ε_1 与 ε_2 的函数，因此我们将以局中人的期望盈利进行分析。

假定老板已经选定 w_H 与 w_L，先考虑第二阶段博弈的 Nash 均衡解 (e_1^*, e_2^*)，对任意 $i(=1,2)$，在给定 $e_j^*(j\neq i)$ 下，工人 i 的期望盈利应为

$$
\begin{aligned}
& w_H P\{y_i(e_i)>y_j(e_j^*)\}+w_L P\{y_i(e_i)\leqslant y_j(e_j^*)\}-g(e_i) \\
&=(w_H-w_L)P\{y_i(e_i)>y_j(e_j^*)\}+w_L-g(e_i) \qquad (5.27)
\end{aligned}
$$

e_i^* 必须如此选取，使得(5.27)式极大化。注意到 $y(e_i)=e_i+\varepsilon_i(i=1,2)$，由一阶条件（即使得导数为零，这里由于 $g(e)$ 为凸函数的假设，其导函数存在应不成问题）得

$$
(w_H-w_L)\frac{\partial}{\partial e_i}P\{y_i(e_i)>y_j(e_j^*)\}=g'(e_i) \qquad (5.28)
$$

(5.28)式指出，工人 i 选取 e_i，使得额外努力的边际负效用 $g'(e_i)$ 等于额外努力的边际利益，或者说等于赢得竞赛的工资获益 w_H-w_L，与赢取概率的边际增量的乘积。

计算

$$
\begin{aligned}
P\{y_i(e_i)>y_j(e_j^*)\} &= P\{\varepsilon_i>e_j^*-\varepsilon_j-e_i\} \\
&=\int_{\varepsilon_j}P\{\varepsilon_i>e_j^*+\varepsilon_j-e_i\mid\varepsilon_j\}f(\varepsilon_j)d\varepsilon_j \\
&=\int_{\varepsilon_j}P\{\varepsilon_i>e_j^*+\varepsilon_j-e_i\}f(\varepsilon_j)d\varepsilon_j
\end{aligned}
$$

$$=\int_{\epsilon_j}[1-F(e_j^*+\epsilon_j-e_i)]f(\epsilon_j)d\epsilon_j$$

(5.29)

(5.29)式中第二个等号是因为概率论熟知的全概率公式，第三个等号是因为 $\epsilon_i,\epsilon_j(i\neq j)$ 互为独立的假设。由(5.29)式，我们可得一阶条件(5.28)式其实为如下形式：

$$(w_H-w_L)\int_{\epsilon_j}f(e_j^*-e_i+\epsilon_j)f(\epsilon_j)d\epsilon_j=g'(e_i) \qquad (5.30)$$

如果 $e_1^*=e_2^*=e^*$，即 Nash 均衡为对称，此时 e^* 满足

$$(w_H-w_L)\int_{\epsilon_j}f^2(\epsilon_j)d\epsilon_j=g'(e_i^*) \qquad (5.31)$$

$g(e)$的凸性，蕴含着 $g'(e)$是 e 的递增函数。因此，如果 w_H-w_L 较大，即竞赛优胜者将获得较多奖励，那么(5.31)式指出 e 必定随之增大，就是说工人必须更努力，这一点似乎非常直观，另一方面，面定奖励不变，即 w_H-w_L 为常数，当产值受到相当干扰时，努力工作显然并不值得，因为此时竞赛的结果很可能由运气而不是由努力来取决。以 ϵ 是均值为零、方差为 σ^2 的正态变量为例作解释，$f(\epsilon)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{\epsilon^2}{2\sigma^2}}$；故

$$\int_{\epsilon_j}f^2(\epsilon_j)d\epsilon_j=\frac{1}{2\sigma\sqrt{\pi}}$$

(5.31)式可以简记为

$$g'(e^*)=c/\sigma \qquad \text{当 } c \text{ 为某常数} \qquad (5.32)$$

可见 e^* 随 σ 增大而减小。σ 作为分布的标准差，刻画了随机干扰的离散程度，它的增大意味着随机干扰变化范围的扩大，于是产值受到随机干扰的影响也增大，干扰的影响掩盖了努力的成果，引起了“侥幸”取胜可能性的上升。

现在向博弈的第一阶段后退。假定如果工人同意参加竞赛(而不是接受其他职业)，他们将通过取(5.31)式所刻画的对称 Nash

均衡解对工资 w_H、w_L 作出反应(这样我们略去了非对称均衡可能性和工人的努力水平选择为隅角解 $e_1=e_2=0$ 而不是由一阶条件求解的可能性);同时假定工人另谋职业的机会将提供赢利 U_a,由于在对称的 Nash 均衡中,每一个工人赢的可能性为 $\frac{1}{2}(P\{y_i(e^*)>y_j(e^*)\}=\frac{1}{2})$。如果老板打算激励工人参加竞赛,他所选择的工资必须满足

$$\frac{1}{2}w_H+\frac{1}{2}w_L-g(e^*)\geqslant U_a \tag{5.33}$$

即参加竞赛所获得的期望赢利不小于 U_a。假定 U_a 足够低,使得老板愿意诱导工人参加竞赛,那么他应在(5.33)的约束之下选择工资以极大化自己的期望赢利 $2e^*-w_H-w_L$。显然,为极大化 $2e^*-w_H-w_L$,最佳之处应当是 e^* 充分地大而 w_H 与 w_L 尽可能地小,根据$g(e^*)$是e^*的递增函数这一假设,我们希望 w_H、w_L 不能过小而$g(e^*)$不能过大,否则将与(5.33)式发生冲突。因此不难明白,在老板选择最佳策略时,必有

$$\frac{1}{2}w_H^*+\frac{1}{2}w_L^*-g(e^*)=U_a$$

或

$$w_L^*=2U_a+2g(e^*)-w_H^* \tag{5.34}$$

于是,老板的期望盈利在最佳策略处应当为 $2e^*-2U_a-2g(e^*)$。因而,老板希望如此地选择工资:使得由此导致的努力水平 e^* 一定满足一阶条件 $g'(e^*)=1$。该条件代入(5.31)式得

$$(w_H-w_L)\int_{\varepsilon_j}f^2(\varepsilon_j)d\varepsilon_j=1 \tag{5.35}$$

故最佳奖励与 w_H+w_L(由(5.34)式可得)分别为

$$w_H^*-w_L^*=\frac{1}{\int_{\varepsilon_j}f^2(\varepsilon_j)d\varepsilon_j}$$

$$w_H^*+w_L^*=2U_a+2g(e^*)$$

这两个方程加上 $g'(e^*)=1$，在 g 为已知，U_a 及 f 均为已知时，可以分别求得解 e^*、w_H^* 与 w_L^*。

§5.5 消耗战与占先博弈

1. 消耗战(War of Attrition)

消耗战实际上是指消耗时间的一类博弈，例如动物的炫耀行为，胜利往往属于炫耀时间最长的一方。两个动物竞争的目的可能是占有某物——比如占有某地或占有某雌性动物，等等。博弈的结果常常是谁坚持到最后谁赢得胜利，当然为此付出的花费也就最多。

本小节中，我们主要讨论一种特殊的平稳消耗战(Stationary War of Attrition)。

先考虑平稳消耗战的离散时间型式，时间呈离散型式：$t=0,1,2,\cdots$。两动物争夺的某物在任一时间 t 具有当时价值 $v(v>1)$；但在博弈过程中，每坚持一个回合它们都将付出 1 的花费。每只动物在每一回合 t 有两种选择：{停止，不停止}。整个博弈过程中，我们最感兴趣的是这样的子博弈：在某时刻 t 至少有一方停止争斗。而在此前的一时刻 $t-1$ 双方均未停止争夺。因此 t 实际上表示了博弈结束的时刻，显然双方的盈利可以表示为 t 的函数。

如果以 δ 表示每一回合的折扣因子，如果局中人 i 首先在 t 停止，称 i 为"先行者"，他的盈利应为

$$L_i(t)=-(1+\delta+\cdots+\delta^{t-1})=-\frac{1-\delta^t}{1-\delta} \tag{5.36}$$

作为"后随者"局中人 i 的对手 j 赢得 v，此时他的盈利应为

$$F_j(t)=-(1+\delta+\cdots+\delta^{t-1})+\delta^t v=L_i(t)+\delta^t v \tag{5.37}$$

作为盈利函数，(5.36)式与(5.37)式也可表示为局中人 j 在 t 停止时双方的盈利，因此在讨论中，只要分出胜负，总以 $L(t)$ 表示输

者的盈利，$F(t)$ 表示赢得 v 一方的盈利。如果两动物在 t 同时停止，我们说谁也没有赢得，此时双方盈利均为 $B_1(t)=B_2(t)=L(t)$。

设想当时间间隔充分短时，上述离散型式就成为连续时间型式的消耗战，$L(t)$与 $F(t)$大致可以描述如图 5.8。

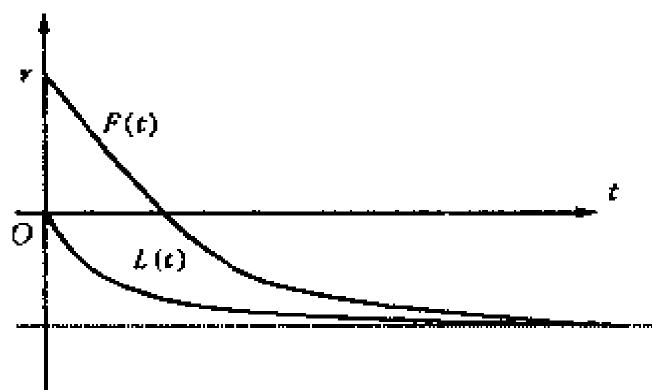


图 5.8

其实很容易找到该博弈的非对称 Nash 均衡：一方的策略是“决不停止”，另一方的策略为“总是停止”。(当然，当某方在 t 已经停止时，就博弈本身而言，另一方再在下一时刻在{停止，不停止}之间进行选择就显得毫无意义，因为此时博弈业已结束。)这种非对称 Nash 均衡总是一方赢一方输，且输者损失最小而赢者花费最小。读者不难发现，到目前为止介绍的消耗战，很有点像我们在完全信息静态博弈中讲述的“懦夫博弈”，消耗战有点相当于两人在独木桥上“顶牛”，谁先退却则为懦夫，而不退却者赢得先过桥。“弱者”总是让“强者”先过桥是一个 Nash 均衡。

存在一个平稳的且包含了混合策略的唯一对称均衡：对任意 p，令“总是 p”表示行为策略“如果其他局中人在 t 之前尚未停止，那么以概率 p 在 t 停止”。等价的策略型混合策略则将概率 $p(1-p)^t$ 赋予纯策略“在 t 停止，如果其他局中人在此之前还没有停止的话”。欲使稳定不变的对称剖面(总是 p，总是 p)成为均衡，对任意 t，在对手先前没有停止的条件下，在 t 停止的盈利，即 $L(t)$，必

须等于

$$p[F(t)]+(1-p)[L(t+1)]$$

即除非对手今天就退出，否则逗留到 $t+1$，然后退出时所获得的盈利。为什么这两种盈利相等构成所需要的必要条件呢？这有些像完全信息静态博弈中猜谜游戏里求混合策略解一样的原理，选择纯策略空间上的概率分布，使得伸出一根手指与伸出两根手指平均来说取得相同的盈利。这里体现为，在 t 之前如果对手不退出的话，那么策略“在 t 停止”与“在 $t+1$ 停止”产生同样的效用。于是解

$$p[L(t)+\delta^t v]+(1-p)L(t+1)=L(t)$$

$$p\delta^t v-p\,\frac{1-\delta^t}{1-\delta}-(1-p)\frac{1-\delta^{t+1}}{1-\delta}=-\frac{1-\delta^t}{1-\delta} \qquad (5.38)$$

立得 $p^*=1/(1+v)$，当 v 从 0 变化直至无穷时，P^* 的范围由 1 向 0 变化。得到这个结论的另外一种方法是，通过再逗留一个周期，局中人以概率 p 得到 v 和以概率 $(1-p)$ 在那个周期损失争斗费用 1，对于该局中人来说，他无所谓是再逗留一个周期还是现在就停止，这必定是这样的情况 $pv=1-p$，于是我们又可得到 $p^*=1/(1+v)$。

通过分析求解，易知 p^* 实际上是唯一的平稳且对称的均衡解的候选者。为了查验它的确是平稳均衡解，仅需注意到，如果局中人 1 取策略“总是 p^*”，那么局中人 2 在每一个可能停止时间 t 时的盈利为 0，读者不妨自己验证一下这个期望盈利。

眼下，读者也许想知道该 Nash 均衡是否子博弈完美的。倘若当局中人想退出的时候，他们可以自由地退出，且不承诺要遵守他们有关停止时间的初始选择，他们想偏离吗？答案是否定的：所有平稳的 Nash 均衡（即，策略独立于日程时间的均衡，我们这里的 Nash 均衡解 p^* 与 t 无关，故称为平稳 Nash 均衡）是子博弈完美。为明白这一点，仅需注意到，盈利的平稳性（即盈利与 t 无关）蕴含

着两个局中人仍很积极争斗的所有子博弈是同类型的，因此，博弈的 Nash 均衡（总是 p^*，总是 p^*）在每一个可能的子博弈中均构成 Nash 均衡。

现在我们需要研究消耗战的连续时间形式。模型仍如上面所述，只不过时间 t 是连续的。在定义盈利函数时，我们需要考虑在时间 t 的折扣因素，也就是在 t 时刻 1 元钱的现时价值。仍假设 1 标准时间的利率为 r，如果我们将标准时间等分为 $1/\Delta$ 个时段，每段长 Δ，且 Δ 很小，近似地将每时段利息视作 $r\Delta$，第一个小时段的折扣因子显然为 $\delta=1/(1+r\Delta)$。时间 t 共有 t/Δ 个时段，此时从离散的时间角度计算：$\delta^t=\left(\frac{1}{1+r\Delta}\right)^{\frac{t}{\Delta}}$考虑到模型为连续时间型，令 $\Delta\to 0$，此时

$$
\begin{aligned}
\delta^t &= \left(\frac{1}{1+r\Delta}\right)^{\frac{t}{\Delta}} \\
&= \left(1-\frac{r\Delta}{1+r\Delta}\right)^{\frac{t}{\Delta}} \\
&= \left(1-\frac{r\Delta}{1+r\Delta}\right)^{\frac{-1}{r\Delta}(-rt)} \to e^{-rt} \qquad (\text{当 } \Delta\to 0 \text{ 时}) \qquad (5.39)
\end{aligned}
$$

因此，在连续时间型消耗战中，原先离散时间的 δ^t 将用 e^{-rt}来代替。现在令 $G_i(t)$表示局中人 i 在 t 或 t 之前停止的概率，$G_i(t)$实际上是累积分布函数。与消耗战的离散时间形式一样，存在一个平稳且对称的均衡 G，具有如下性质：在每一时刻 t 局中人关于下面两种策略取哪一个为好是无所谓的：①在时刻 t 停止；②再稍微等一会儿，等到 $t+\varepsilon$（ε 为相当小的任意正数），看看对手是否先停止。

基于在 t 之前尚未停止的条件（该事件的概率为 $1-G(t)$），局中人再等 ε 时间期望获得 v 的概率为 $\varepsilon \mathrm{d}G(t)/(1-G(t))$，而为此多付出的损失应近似地为$\int_0^{\varepsilon} e^{-rt}\,\mathrm{d}t=\frac{1}{r}(1-e^{-r\varepsilon})\approx\varepsilon$。如果局中人在 t 时对于停止或再稍等一会儿表现出无所谓，那么意味着稍等

一会儿的期望盈利等于零：

$$\frac{v\cdot\varepsilon\mathrm{d}G(t)}{1-G(t)}-\varepsilon=0 \tag{5.40}$$

即

$$\frac{\mathrm{d}G(t)}{1-G(t)}=\frac{1}{v} \tag{5.41}$$

因此，G 即为负指数分布 $G^{*}(t)=1-\exp(-t/v)$。根据上述分析，读者不难验证，如果局中人 1 采用策略 G^{*}，那么局中人 2 在任何停止时刻的期望盈利为零，故 $G^{*}(t)$ 为平稳对称均衡。至于它是否子博弈完美，仅需注意到均衡的平稳性即可。

有趣且有意义的是，连续时间型消耗战的对称均衡可以证明是离散时间型消耗战对称均衡的极限。这些离散型模型实际上是将 1 标准（或实际）时间划分为 $1/\Delta$ 个长度为 Δ 的小时段，如前面讨论折扣因素时一样。这样在每个时段内的争斗花费为 Δ。争夺目标 v 值不随时段变化而变化，因为在离散型与连续型中我们取 v 值均为"硬件"。对于这样的离散时间型博奕，均衡策略由求解 $p^{*}v(1-p^{*})\Delta=0$ 而得，因此 $p^{*}=\Delta/(\Delta+v)$。对于给定的实际时间 t，在 0 至 t 之间存在 $n=t/\Delta$（可以取适当的 Δ 使 t/Δ 为整数，反正 Δ 随后将趋于零）个时段。在离散型模型的公式中，局中人在 t 之前不停止的概率应为

$$\begin{aligned}1-G(t)&=(1-p^{*})^{n}\\&=\left(\frac{v}{v+\Delta}\right)^{\frac{t}{\Delta}}=\left(1-\frac{\Delta}{v+\Delta}\right)^{\left(-\frac{v+\Delta}{\Delta}\right)\cdot\left(-\frac{t}{v+\Delta}\right)}\\&\to e^{-\frac{t}{v}}\qquad(\text{当 }\Delta\to 0)\end{aligned} \tag{5.42}$$

于是，当 $\Delta\to 0$ 时，对称离散时间均衡收敛于对称连续时间均衡。

连续时间型消耗战的一个有趣且著名的例子由 Parker 于 1970 年提供。假设有一群雄蝇叮在奶油渣上，等待雌蝇飞来再进行交配。通常，雌蝇更喜欢新鲜奶油渣，因此，雄蝇在原先停留的奶

油渣上逗留一会儿后会去寻找更新鲜的奶油渣。这是一个多蝇消耗战博弈，但与2蝇博弈基本原理一致：假如大多数雄蝇早早地离开奶油渣，那么留下的雄蝇就可"占有"后来飞来的雌蝇进行交配；反过来，如果大多数雄蝇长时间停留在一块奶油渣上，那么另找新鲜的奶油渣是雄蝇的较优选择。Parker测定了蝇叮在一块奶油渣上的时间，大量数据表明，它符合负指数分布这一规律。

2. 占先博弈(Preemption Games)

消耗战相当于持久战，谁坚持到最后，尽管付出一定代价，但最终微笑的人就是他。然而，在经济领域内并不总是以持久战形式进行博弈。譬如，开辟新的市场，采纳新的技术等等，通常有"先下手为强"的优势。好比桌子上放着1元钱的纸币，有两个局中人在t时刻去"抓"，谁先抓则谁得钱，这个典型的占有博弈称为"抓钱博弈"(grab the dollar)。对于占先博弈，我们也可以定义时刻t的$L(t)$与$F(t)$，以及$B(t)$。$L(t)$为先抓者在t时刻的盈利，$F(t)$为后抓者在t时刻的盈利，与消耗战不同的是，对某些t来说，这里成立$L(t)>F(t)$。至于$B(t)$，则表示两人同时"抓"的盈利。为了与消耗战作比较，在许多文献中常把这里的"抓"称作"停止"，希望读者对此不要产生混淆。

我们展开对"抢钱博弈"的讨论。时间t为离散形式：$t=0,1,2,\cdots$。桌上放有1元纸币，两个局中人，可以两个人一起或者任何一个人试图去抓。如果只有一个人去抓，那个局中人获益1(元)而另一个为0；如果两个人同时抓，纸币被损坏，游戏规定每人罚出1(元)；如果没有一个去抓，1元钱仍静静地躺在桌上，两个局中人均无所收获。设每个局中人的折扣因子都取δ，这样，在每一时刻t，应有$L(t)=\delta^t,F(t)=0,B(t)=-\delta^t$。该博弈有显然的非对称均衡：某个局中人以概率1"赢"。

现在我们考虑抓钱博弈的对称混合策略均衡，其中，在每一个时刻t，每一个局中人以1/2概率抓到1元钱。设想局中人i，在t

时刻如果不抓的话，其盈利必定为 0；如果去抓，设他抓到 1 元钱的概率为 $p(t)$，那么他面临两种情况，倘若在 t 之前局中人 j 已经将钱抓走，此时 i 的盈利只能为 0，而若局中人 j 在 t 之前尚未抓走钱，那么局中人 i 以概率 $p(t)$ 得到 1 而以 $(1-p(t))$ 概率受罚 1 元钱，考虑到折扣因子，此时他的期望盈利为 $\delta^t[p(t)-(1-p(t))]$。所谓混合策略均衡，要求局中人 i 在每个 t 时，在抓与不抓之间感到无所谓，因此必有 $\delta^t[p(t)-(1-p(t))]=0$，即 $p(t)=1/2$。对于局中人 j 也可同样考虑，因此，在对称的混合策略均衡中，盈利向量应为(0,0)。我们知道，动态博弈的混合策略需要在各个结局上的概率分布。先不妨考虑单局期情况，即 $t=0$ 时，显然 p(局中人 1 先单独抓) $=p$(局中人 2 先单独抓) $=p$(局中人 1 与 2 同时抓) $=p$(局中人 1 与 2 都不抓) $=1/4$。现讨论在任何 t 时刻，局中人 1 在该时刻先单独抓得 1 元钱这个事件是建立在此之前 t 个时刻要么两人都不抓，要么两人都抓而被罚(从而桌上仍放有 1 元钱)的基础上，由于各时期的博弈可以视作独立的，因此

$$P\{\text{局中人 1 在 } t \text{ 时刻先单独抓得钱}\}=(1/4)^{t+1}$$

同样

$$P\{\text{局中人 2 在 } t \text{ 时刻先单独抓得钱}\}=(1/4)^{t+1}$$

在 t 时刻两个人同时抓这一事件也是建立在上述前提上，因此

$$P\{\text{两个局中人在 } t \text{ 时刻同时抓钱}\}=(1/4)^{t+1}$$

一个有趣的事实是，尽管我们在动态过程中考虑到了每个 t 时刻的折扣因子 δ，但是混合策略均衡所涉及的概率均与 δ 无关，于是这些概率自然与周期长度 Δ 无关。周期长度 Δ 的作用在消耗战研究中已经提到，在那里利用混合概率与 Δ 成比例的特点而引导我们导求连续型消耗战的均衡分布。现在，抓钱博弈在这一点上与消耗战恰好不同，但是，这个有趣的事实却使得我们比较扫兴，因为它使得我们在导求博弈的连续时间表示型式方面遇到了困难。

读者也许会问，困难的症结何在？因为在通常情况下，我们常

将连续情况分割成离散的格子，从得到的结果中再令格子长度$\Delta\to0$时取极限。但是我们稍稍地讨论一下这种办法立即可以觉察利用这种办法也许暂时无法得到“抓钱博弈”的连续时间形式表示。

先暂时令t为固定的正数，不妨设Δ是可以使t/Δ成为正整数$n=t/\Delta$的小区间长度，将区间$[0,t]$分割为n个区间，现在我们考虑当$\Delta\to0$时，我们在前面离散情况的均衡分布将会发生怎样的变化。显然在t时刻之前，至少有一人抓钱的概率可近似地视作$1-\left(\frac{1}{4}\right)^n=1-\left(\frac{1}{4}\right)^{\frac{t}{\Delta}}$，当$\Delta$趋于0时，该概率趋于1。由于抓钱博弈的混合策略均衡概率不依赖于δ，所以在这里干脆不管δ有多大。局中人1在t之前赢得1元钱的均衡分布概率近似地为

$$\frac{1}{4}+\left(\frac{1}{4}\right)^2+\cdots+\left(\frac{1}{4}\right)^{\frac{t}{\Delta}+1}=\frac{1}{4}\left[1+\frac{1}{4}+\cdots+\left(\frac{1}{4}\right)^n\right]$$

$$=\frac{1}{4}\cdot\frac{1-\left(\frac{1}{4}\right)^{n+1}}{1-\frac{1}{4}}\to\frac{1}{3}\quad(\text{当 }\Delta\to0\text{ 时})\tag{5.43}$$

注意到(5.43)式取极限时仅需t不等于0，但对任何正数t均成立，因此在(5.43)式中若再令$t\to0$，我们得到的极限仍为1/3。这表明，局中人1在$t=0$时抓到钱的概率为1/3。同样的讨论可以得出，局中人2在$t=0$时抓到钱的概率也为1/3；两个局中人在$t=0$时同时抓的概率又为1/3；还可得到，博弈在$t>0$继续进行的概率为0。这个结果将使我们大吃一惊，它与我们在前固分析到的在$t=0$时一次性抓钱博弈的均衡策略完全不同。看来，这个极限分布不能表示为迄今为止我们讨论的连续时间策略型的均衡。因为如果博弈在$t=0$时以概率1结束，那么，对于至少有一个局中人i，其在$t=0$处的期望盈利应等于1，而它的对手得到钱的概率为0，这不是一个对称的均衡。问题在于不同的离散时间策略剖面序列收

敛于一个极限：博弈概率为 1 地在出发处结束。Fudenberg 与 Tirole 于 1985 年对他们自己研究分析的特殊的占先博弈提出了推广形式的连续时间策略与盈利函数以作为离散时间均衡的极限。Simon于 1988 年将该方法推广到更宽广的一类博弈中去，我们在本书中不准备作详细介绍。

占先博弈的第二个例子，我们假定 $L(t)=14-(t-7)^2$，$F(t)=0$，以及 $B(t)<0$。这些盈利函数描述了如下情况：两个公司中的任一个可以推出新产品，该产品不影响它们已有的商务，因而 $F(t)=0$，固有成本加上过分的双寡垄断价格一起蕴含着，如果双方开发新产品将使两个公司均有所损失。于是，一旦某公司推行新产品，另一个公司决不再这样做，除非他们恰好同时开发新产品。为了避免必须考虑此类失误和与此相连的混合策略的可能性，可以作简单的假设，局中人 1 仅在 $t=0,2,\cdots$ 这些偶数周期内停止，而局中人 2 仅在 $t=1,3,\cdots$ 这样的奇数周期内停止，使得博弈成为具有完美信息。从而局中人有 3 个 Pareto-有效结局，要么局中人 1 在 $t=6$ 或 $t=8$ 时停止，要么局中人 2 在 $t=7$ 时停止。相应的 Pareto-有效盈利向量为(13,0)和(0,14)(这 3 个有效结局很容易得到，读者只要注意到 $L(t)$ 的表达式立即可知这个事实)。现在来考虑该占先博弈的子博弈完美均衡，占先博弈要求 $L(t)>F(t)$，要不谁也不愿占先，注意到时间 t 为离散型，$L(0)=-35<0$，$L(1)=-22<0$，…，$L(3)=-2<0$，$L(4)=5>0=F(4)$，对于公司 1 来说，它应当毫不犹豫地在 $t=4$ 时停止，一旦它偏离该策略，那么在 $t=5$ 时它将可能仅获盈利 0。因此，公司 1 在 $t=4$ 时停止是博弈的唯一子博弈完美均衡，此时的均衡盈利为(5,0)，小于 Pareto-有效盈利(13,0)。这是一个"浪费经济收益"(rent dissipation)的典型例子：尽管在以后引进(或开发)新产品将得到可能更多的经济收益，但在均衡状态，第一次竞争强使在经济收益为非负时的第一时刻开发新产品。

§5.6 开环和闭环均衡

在多阶段博弈中存在两种不同的信息结构，我们将用开环(open-loop)和闭环(closed-loop)这两个专用名词分别命名各自的信息结构；相应的策略、均衡则也在前面冠上开环和闭环这两个名词。因此在理论上，开环均衡和闭环均衡不应当视作新的均衡概念，它们实质上是描述不同信息结构的特殊一类博弈均衡的手段与方法。

有些博弈，局中人除了自己的行动与日程之外看不到任何历史，或者在博弈的一开始局中人必须选择仅依赖于日程时间的行动日程表。例如，两阶段猜谜游戏中，局中人只要确定自己在第一次及第二次时各伸出几根手指，(a_1, a_2)就是局中人 1 选择的行动日程表：在第一次猜谜中伸出 a_1 根手指，在第二次猜谜时伸出 a_2 根手指。这些博弈的策略特点为，它们只是日程时间的函数，所有这类博弈的策略称之为开环策略，所有的 Nash 均衡都在开环策略之中。以开环策略构成的均衡称为开环均衡。在两阶段二人猜谜游戏中，混合策略{((1/2,1/2),(1/2,1/2)),((1/2,1/2),(1/2,1/2))}是开环均衡。

相当多的情况，博弈局中人在选择自己的行动时需要根据自己所看到的历史，尤其是对手在此之前采取的行动而作出决策，这类博弈的信息结构就是闭环信息结构。Stackelberg 竞争模型中，后随者将在先行者选择产量的基础上选择使自己盈利达到极大化的产量，就是一个典型的例子。展开一下，可观察行动的多阶段博弈都属于这种情况。在这一类博弈中，局中人的策略不仅依赖于日程时间，还依赖于其他的变量。为了对由自然产生的外生行动作出反应，为了对由对手实行的混合策略作出反应，为了对由于对手可能偏离均衡策略作出反应，局中人可能不愿意让自己的策略仅依

赖于日程时间，或者说，局中人可能不愿意使用开环策略，他们愿意使用闭环策略。

如果在一个展开型博弈问题中，既有开环策略，又存在闭环策略，那么子博弈完美均衡通常不会是开环的。这个问题从直观上很容易理解，因为这个展开型博弈的信息结构必定是闭环式的，采用开环策略，就是抛开已观察到的信息，完全“机械地”按照既定的日程行动，显然很难达到极大化自己的盈利这一目的。在实际博弈中，通常局中人总是充分利用自己获得的信息（若可能的话）以进行理性行动。这种直观理解换成子博弈完美的“正统”术语来讲，子博弈完美要求局中人对于随机的行动以及对于未料到的偏离作出最佳反应；特别地，迎合这个条件的开环策略则要求，无论对手在过去是否偏离，采取同样的行动是最优的办法。这在通常情况下可能性很小。术语闭环均衡通常是指，局中人可以观察到其对手在每个周期结束时的行动并对此作出反应的博弈中的子博弈完美均衡。当然，拥有此类信息结构的博弈可以存在不完美的 Nash 均衡。我们仍以 Stackelberg 竞争模型为例，如前所知，Stackelberg 均衡是子博弈完美，后随者的行动是观察到先行者行动后再作决策的，它当然也是闭环均衡。但是这个博弈属于“决策论”。后随者的决策可以只依赖时间，回忆一下有关 Stackelberg 模型的分析，后随者可以不管先行者的产量是多少，它总是取 $q_2=4$，以迫使先行者也取 $q_1=4$，这是一个纯策略 Nash 均衡，正如我们已经指出过，它依赖于一个空头威胁，因此这个开环策略剖面尽管是 Nash 均衡，但是它不是完美的。

通常，刻画一个给定情况的开环均衡比起刻画闭环均衡来要容易得多。在某种程度上，是因为闭环策略空间相当地大。这就是在经济分析中采用开环均衡的一种解释。对开环均衡感兴趣的第二个原因是，它们可以作为讨论闭环信息结构中策略刺激效果的一个有用的基准点，这里所说的策略刺激是指刺激改变当前的行

动以至影响对手将来的行为。至于第三个原因，如果存在许多“小”局中人的话，开环均衡可以是闭环均衡的一个良好近似。直观上，假如局中人很“小”，由对手引起的未预料到的偏离可能对于局中人的最优行动产生微不足道的影响。以上谈到的三种原因，第一条是没有什么值得怀疑的，我们将在下面展开的讨论中详细地讲述第二种原因，至于第三种原因，由于其数学内容较深，我们认为不适合国内经济类读者而只得割爱。

利用开环均衡作为测度策略效果的基准点。

我们将构造一个例子来阐述这个问题，为方便起见，所构造的博弈具有连续的行动空间。

考虑两人二阶段博弈，在第一阶段，局中人 $i(i=1,2)$ 各自从 A_i 中同时选择自己的行动 a_i，在第二阶段，局中人 i 各自从 B_i 中同时选择自己的行动 b_i，A_1、A_2、B_1、B_2 均假设为某实数区间。再设局中人 i 的盈利函数 u_i 为可微函数且关于他自身的行动是凹函数。

设日程路径 (a^*,b^*) 是博弈的开环均衡，其中 $a^*=(a_1^*,a_2^*)=(a_i^*,a_{-i}^*)$，$b^*=(b_1^*,b_2^*)=(b_i^*,b_{-i}^*)$ $(i=1,2)$，作为均衡，(a^*,b^*) 必须满足

$$a_i^* \text{ 使 } u_i((a_i,a_{-i}^*),b^*) \text{极大化} \tag{5.44}$$

$$b_i^* \text{ 使 } u_i(a^*,(b_i,b_{-i}^*)) \text{极大化} \tag{5.45}$$

由于我们假设 u_i 关于 a_i 与 b_i 为凹函数，因此通过一阶条件

$$\frac{\partial u_i}{\partial a_i}=0 \text{ 与 } \frac{\partial u_i}{\partial b_i}=0 \tag{5.46}$$

可以得到 A_i、B_i 中的内点解。

现在假定该博弈存在闭环均衡（注：如果我们的模型规定，在第二阶段开始前，各局中人能观察到第一阶段的结局，这样，博弈具有闭环信息结构，于是关于闭环均衡存在的假设就有其合理性）。在这个闭环均衡中，在第一阶段任何行动 a 之后的第二阶段

行动 $b^*(a)$ 应该是第二阶段博弈的 Nash 均衡。就是说，对于每一个 $a=(a_1, a_2)$，$b_i^*(a)$ 使 $u_i(a, b_i, b_{-i}^*(a))$ 达到极大化，由于均衡是闭环的，b^* 是 a 的函数，现假设 b^* 是可微函数。局中人 i 选择最优 a_i 的一阶条件为

$$\frac{\partial u_i}{\partial a_i}+\frac{\partial u_i}{\partial b_{-i}}\cdot\frac{\partial b_{-i}^*}{\partial a_i}=0 \tag{5.47}$$

比较(5.46)式与(5.47)式，局中人 i 为使 u_i 极大化而对 a_i 进行选择的一阶条件在开环与闭环之间的差异是闭环一阶条件多了一项 $\frac{\partial u_i}{\partial b_{-i}}\cdot\frac{\partial b_{-i}^*}{\partial a_i}$。正是这一项对应于局中人 i 改变 a_i 以影响 b_{-i} 的"策略刺激"。为了对这一点有个具体理解，我们用一个两阶段 Cournot 竞争模型进行解释，该模型中行动选择是产量。设想在生产过程中存在着"通过实践而学习"的现象，因此每个公司若在第一阶段生产得多的话，那么它在第二阶段生产的边际成本就会相应减少，或者说，第二阶段的边际成本是第一阶段产量的递减函数。一旦给定第一阶段两个公司的产量，相当于给定它们在第二阶段生产的边际成本，那么第二阶段的 Nash 均衡 $b^*(a)$ 实际上就是 Cournot 均衡。不妨先考虑第二阶段 Cournot 竞争，设需求函数为 $p(q)=1-q=1-(b_1+b_2)$，公司 i 的边际成本为 $c(i=1,2)$，那么 Cournot 均衡产量分别为

$$b_1^*=(1-2c_1+c_2)/3 \quad b_2^*=(1-2c_2+c_1)/3 \tag{5.48}$$

(5.48)式展示了一个事实，公司 i 的 Cournot 均衡产量是其对手的边际成本的递增函数。但在这里，$c_1=c_1(a_1)$，$c_2=c_2(a_2)$ 且 c_i 随 a_i 的增加而减少。因此若增加 a_2，则减少 $c_2(a_2)$，也就降低了 b_1^*，同理，a_1 的增加却使 b_2^* 减少，用数学语言来说，在两阶段 Cournot 竞争博弈中，对于每一个 i，$\frac{\partial b_{-i}^*}{\partial a_i}$ 为负。开环均衡的第一阶段产量为 a^*，那么 $\frac{\partial b_{-i}^*}{\partial a_i}$ 在 a^* 处为负，假如公司 i 希望 b_{-i} 减少的话，蕴含着公

司 i 应增加 a_i 以超过 a_i^*，这就是公司 i 的"策略刺激"。在 Cournot 竞争中。产量的最优选择为 $q_i=(1-q_j-c_i)/2$，每个公司相应的盈利为 $u_i=(1-q_j-c_i)^2/4$，注意到 $q_1+q_2<1$，因此每个公司都希望对手的产量低一些。于是，这个模型中的公司 i 的"策略刺激"赞成第一阶段用于"学习"的追加投资超过公司在开环均衡中选择产量的投资。

如果第二周期均衡行动随第一周期行动的增加而增加的话，且公司也希望其对手的第二周期行动 b_{-i} 低一些，那么，策略刺激倾向于将第一周期行动从开环均衡水平减少下来。

§5.7 有限与无限状态均衡

熟悉微积分的读者一定知晓，倘若一个级数收敛，那么它的一般项随 n 的增加一定趋于零，因此，当 n 相当大的时候，级数与部分和 S_n 几乎可以视作"相同"的。在无限状态展开型博弈中，我们也面临类似的问题。回忆一下我们在讨论无限状态的一步偏离准则时，引进了"在无穷远处连续"的概念，已经指出，它意指在相当大的 t 之后的事件，其影响是微不足道的。因此人们可以期望在这种情况下，"相同博弈"的有限与无限状态的均衡将是紧密相关联的。事情的确是这样，但并不是说所有的无限状态均衡是相应的有限博弈均衡的极限。图 5.1 的重复囚徒窘境就是一个例子。在那里，我们指出，当折扣因子 $\delta>1/2$ 时，"合作"是无限博弈的子博弈完美，但是由于在任何有限囚徒窘境中均衡，"合作"不可能是 Nash 均衡，因此，无限囚徒窘境中至少有一个均衡不是有限博弈的极限。针对这个问题，1980 年 Radner 指出，如果人们放宽局中人不折不扣地极大化自己盈利的假设，那么"合作"可以修改为有限囚徒窘境均衡。Radner 的主要手段是引进了如下新概念：

定义 5.1 一个策略剖面 σ^*，如果对所有局中人 i 和所有策略 σ_i，对某个正数 $\epsilon>0$ 成立

$$u_i(\sigma_i^*,\sigma_{-i}^*)\geqslant u_i(\sigma_i,\sigma_{-i}^*)-\epsilon \qquad (5.49)$$

那么，我们称 σ^* 是一个 ϵ-Nash 均衡。

如果没有一个局中人可以在任何子博弈中通过偏离而使自己获益增多 ϵ，那么这样的策略剖面是 ϵ-完美均衡。

ϵ-Nash 均衡的放宽条件就在于(5.49)左边的 ϵ，这相当于放宽局中人的"理性"要求，在小小的 ϵ 尺度范围内，局中人的"理性"促使他愿意接受现实。放宽 ϵ 尺度并不是数学游戏，它有着一定的现实意义。在现实生活中，局中人对盈利的计算不可能绝对地精确。因此，由于某些不确定性原因，理性的局中人会甘愿忍受微小 ϵ 的"亏本"。

现在考虑效用函数在无穷远处连续的无限博弈 G^∞，对于每一个有限 T，我们构造 T 周期"截断"博弈 G^T：通过选择 G^∞ 中任意策略 $\tilde{s}$，并确定在 T 以后的博弈采取策略 $\tilde{s}$，这样构成的一个无限形式博弈相当于 G^∞ 在 T 处被截断。"截断"博弈的策略 s^T 确定在 T 周期之前(包括 T 周期在内)所有周期的行动，以及在 T 之后的周期内取策略 $\tilde{s}$。为了符号的简化起见，"截断"博弈 G^T 的策略和 G^∞ 内对应部分的策略一起都用 s^T 来表示，当我们谈及截断博弈策略收敛时，将 s^T 看作为 G^∞ 中的策略。于是，利用 G^∞ 中的盈利函数可以导出 G^T 的盈利函数(注：读者千万不要认为"截断"博弈是将无限博弈在有限周期处截断变成完完全全的有限博弈)。尽管我们不准备用严格的数学来进行讨论(这对那些只关心博弈论的经济应用的读者也许有些难度)，但是读者不难设想，由于对于相当大的 T，T 以后的事情仅有微不足道的影响，因此 G^∞ 的均衡可以利用 G^T 的策略剖面的极限来刻画。

仍以图 5.1 的囚徒窘境为例，考虑 $\delta>1/2$。构造"截断"博弈

G^T，在 T 周期之后的所有周期内，两个囚徒均背叛。在 T 个周期的有限囚徒窘境中，一直(背叛，背叛)显然是唯一的子博弈完美均衡。然而考虑 G^T，对于每一个局中人，如果他的对手的策略是合作直到发生背叛为止，且此后一直采取背叛行动，那么他的最佳反应无疑为合作直到 T 周期之前，且在 T 周期背叛并一直背叛下去。计算他的平均效用为

$$\frac{1-\delta}{1-\delta^{T+1}}(1+\delta+\cdots+\delta^{T-1}+2\delta^T+0+0+\cdots)$$
$$=\frac{1-\delta}{1-\delta^{T+1}}[(1+\delta+\cdots+\delta^T)+\delta^T]$$
$$=1+\frac{\delta^T(1-\delta)}{1-\delta^{T+1}} \tag{5.50}$$

当 $T\to\infty$时，(5.50)式显然趋于 1。注意到 G^∞中“永远合作”的盈利为 1。G^T 的这个策略与这个最佳结果之间的差异当 $T\to\infty$时趋于 0，因此不难知道，在 G^T 中，有限 T 个周期内均取合作态度其实是一个 ε 完美均衡。G^∞中的子博弈完美均衡“永远合作”可以用一系列 G^T 的上述策略的极限来刻画。

第六章　讨价还价模型

人们对讨价还价，已经是相当熟悉的了，因为这可以说是经济生活中的极重要部分。从日常的货物买卖到国际贸易乃至重大政治问题的谈判，都存在着讨价还价。最简单的二人买卖，双方轮流提出自己的开价，实际上就是一个可观察行动的多阶段博弈，我们不把它放入上一章而单独列为一章，完全是由于该模型在经济博弈论方面的重要地位。

§6.1　不存在耐心问题的讨价还价

两人为买卖一物谈判一个价格。买者 B 心目中愿意最高出 300 元买此物，卖者 S 将不接受任何低于 200 元以下的开价。注意，在他们各自最高与最低价之间只有存在正差异才有可能进行谈判并继续下去。在我们的模型中，假定以上这些是双方的共同知识。买者愿付的最高价格 300 元是买者 B 的保留价，而卖者愿接受的最低价则为卖者 S 的保留价。谈判价与各自的保留价之间的差恰为各人从交易中获得的盈利，显然，B 与 S 的获益范围从 0 至 100 元，我们可以将这 100 元视作双方欲分的蛋糕。这对以后建立讨价还价模型并进行讨论分析是很方便的。

很容易用一个动态模型来描述 B 与 S 之间的价格博弈：B 先提出一个开价，S 有两种选择，如果 S 接受则博弈结束，如果 S 拒绝，那么他然后提出自己的开价，如果 B 接受则买卖成交，如果 B

不接受，则买卖不成交而结束博弈（其实还存在继续讨价还价下去的模型）。其展开型如图 6.1 所示（其中 a 表示接受，R 表示拒绝）。

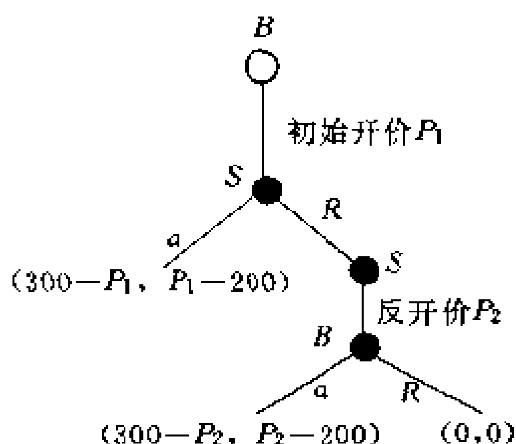


图6.1

图 6.1 中，“初始开价”、“原开价”的选择范围为连续区间。终点结的盈利是双方从交易中的获益，倘若一个局中人在接受或拒绝对方开价方面表现出无所谓，则假定为接受开价。从博弈树来看，博弈的最先行动者与最后行动者都是 B，他的最初开价 P_1 将使 S 对此作出决策：成交或者反过来提出开价 P_2，然而 B 是否接受 P_2 又决定了博弈的结果。但是，如果从“开价”角度考虑，B 是先行者，由他提出初始价格，S 是最后行动者，因为由他最后开价。讨价还价模型的关键是价格，因此在这里通常认为 S（卖者）具有最后的“实际”行动。本博弈存在两个轮次（或回合）。每一个轮次，由一方开价，另一方作出“接受”或者“拒绝”决策。轮次少，时间短，从而不存在成交价格的现时价值问题，于是这样的讨价还价被称为不存在有无耐心的问题。

由图 6.1，我们可以使用后退归纳原理。只要 S 最后的要价不超过 300 元，买者 B 将会接受。例如 P_2=299 元，盈利向量(1,99)优于(0,0)，这就是 B 接受 P_2 的简单原因（当然，P_2 可以取作 300−ϵ 元，ϵ 为任意小的正数，从理论上说 B 也将接受。现实生活意义使我们不去考虑像 ϵ 那样的微利）。沿博弈树回到讨价还价的第一

轮次，显然，S 将拒绝由 B 提出的任何小于 300 的开价（除非 B 开出的价正是他在下一轮次将提出来的 $P_2=299$（或 $300-\varepsilon$）），因为在下一轮次 S 站在一个非常有利的地位。分析结果表明，B 在博弈的一开始就以几乎 300 元的价格向 S 买下自己需要的东西，因为他明白不管怎样，他将付出近 300 元来结束买卖。这个讨价还价模型的特点是，卖者 S 作为最后开价者，他享受着“后动者优势”，这一优势使得他足以“吃掉几乎整块蛋糕”。

倘若设想讨价还价模型由 B 开价只进行一次，然后要么成交要么买卖不成而结束博弈。其展开型表示为图 6.2。

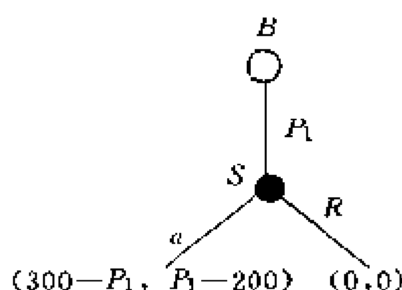


图6.2

此时，卖者 S 没有机会展现他开出第二个有利于自己的价格的优势，从盈利角度分析，只要 B 开出的价格不低于 200 元，买卖应当成交。其实，根据原先的假设，S 会接受 200 元，因为他对于“接受”或“拒绝”200 元这个价格无所谓（两者盈利均为 0），这在我们的模型里蕴含着 S 接受这个价格。在实际生活中，B 只要开出略高于 200 元的价钱就足够了。

因此，我们发现，此类讨价还价模型的预测结果与两个因素有密切联系：谁是最后开价者和开价的轮次数目。这与日常生活中所见非常吻合，买者 B 很想买下自己急需的东西，则卖者 S 将处于一个有利的地位；反过来，如果卖者 S 急于推销自己的商品，则买者 B 会处于一个相对有利的地位。一个有趣但似乎“荒谬”的情况是，根据以上类似的分析，在这个博弈模型中，如果 B 作出初始开价且开价轮次为奇数，那么 B 将“几乎吃掉整块蛋糕”，倘若开价

轮次为偶数的话，S 将"几乎吃掉整块蛋糕"。问题的"推广"使人感到预测结果似乎不太现实，我们将在以后继续进行讨论。

§6.2 无耐心的讨价还价

上一节我们讨论的讨价还价模型中，开价的轮次只有 1 或 2 个。本节将把轮次数较大地扩大，例如讨价还价 100 次。更重要地，假设拖延达成一项协议就会对谈判者强加一定成本或花费。这种成本也可以同该谈判者的将来所获相对于现时价值的折扣联系在一起。当交易者的时间具有较大机会成本时，即使少数几天的延迟谈判，价值也是昂贵的。

仍以上节 B 与 S 二人之间的买卖谈判为例，不过这次将轮次数目增加到 100。我们假定每拖延一轮协议的成本将会使两个局中人都缩减从交易中所获益的 3%。现在使用后退归纳法，考虑第 100 轮时的开价，由于 100 是偶数，根据 §6.1 末的推理，应当 S 吃掉几乎整块蛋糕，从理论上讲，即双方保留价格之间的 100 元盈余全由 S 分得。虽然我们认为画出 100 轮讨价还价博弈树几乎是"荒谬"的，还是可以设想后退到第 99 轮由 B 提出开价。由于一轮延迟将使 S 付出 $100\times3\%=3$ 元的成本，因此 B 提出 297 元的开价并相信 S 会接受。因为对 S 来说，虽然等到下一轮由他开价，可以获得全部 100 元盈余，但是由于他的"无耐心"，他将乐意在这一轮接受 97 元盈余的方案，这一轮的盈利 97 元与下一轮的实际盈利 $=100-3$(延迟一轮的成本)$=97$ 元完全一样，S 对这一轮是否接受将感到无所谓，根据模型假设，S 会接受 97 元盈利的建议。此时 B 将获得 $100\times3\%=3$ 元的盈利。考虑再后退到第 98 轮由 S 开价，S 应考虑到，因为 B 在第 99 轮接受 3 元盈利，B 应当愿意在第 98 轮接受比 3 元少 3%($3\times0.03=0.09$ 元)的盈利，也即 B 接受 2.91 元盈利。于是 S 自然开价 297.09 元而使自己获得盈利 97.09

元(即 100 元的 97.09%)。这不仅比让他再等一轮可多得 0.09 元,而且这样的开价使得他节省了等待的成本。再一次后退到第 97 轮由 B 开价。因为 B 知道 S 无耐心,由于 S 在第 98 轮接受 97.09 元盈余,因此 B 认为 S 应当愿意在本轮次接受比 97.09 少 3%的盈余,即少 97.09×0.03=2.91 元,于是 B 开价为 294.18 元,B 自己将接受 5.82 元盈余,大于他再等一轮的获益。

我们已经看到,对买卖两者来说,每一轮次的谈判(包括一方开价,另一方表示是否接受总共两个阶段)都强加了成本。只要通过对另一局中人有足够吸引力的盈余分配,该成本是可以避免的。一直持续后退归纳法直到第一轮由 B 开价,那么最后三轮的各自的最优开价可见表 6.1。

表 6.1

轮　次	开价者	S 享有盈余	B 享有盈余
3	B	51.80	48.20
2	S	53.25	46.75
1	B	51.65	48.35

我们最终发现,B 的最优初始开价为 251.65 元,S 将会接受这个开价。B 将以自己获盈利 48.35 元而 S 获盈利 51.65 元来结束双方之间的讨价还价。注意到盈余的分配几乎是一半对一半,当然,S 作为最后一个开价者,继续享有少许后动者优势。其实,我们还可以考虑更多的开价轮次,例如,120 甚至 120 以上,我们会发现继续后退的结果将使双方的盈余在 50 元上下摆动。

使用上述分析,盈余的精确分享依赖于开价的轮次以及交易成本的数量。50 对 50 的分享发生在:拖延的成本对双方是相等的(此时我们称讨价还价的双方具有对称的无耐心)、开价轮次数较大和一个周期的拖延成本较小的情况。虽然非正式的经验主义常建议选择一半对一半分享,但并非总是那样。假如推迟成交对卖者

来说比买者花费更多，那么 S 更急于成交从而“更多地无耐心”。例如假定推迟一次将使卖者 S 从交易所得中减少 6%，而买者 B 拖延一次仅减少交易所得的 3%。这种场合，卖者比买者更焦急更缺乏耐心，与上述局中人具有对称无耐心不同，我们称此时的局中人具有不对称的无耐心。在这种情况，设想 B 与 S 之间仍进行 100 轮次开价。仍应用后退归纳法，先考虑第 100 轮次由 S 开价，与对称无耐心情况完全一样，S 的最佳开价是 300 元，获得 100 元盈余中的全部，因为他知道 B 将接受这个开价。后退到第 99 轮次由 B 开价，现在的情况开始与对称无耐心时有所不同，由于再等一轮次 S 将花费 100 元×6%＝6 元，因此 B 的开价应为 294 元(即 B 获盈余的 6%)。而若后退到第 98 轮次，考虑到 B 再等一轮次将花费 6 元×3%＝0.18 元，因此 S 在开价时将他在第 99 轮次获盈余的 94 元再加上 0.18 元，从而开价 294.18 元，…，如此一直后退下去。依上面所述规律，不难用 Microsoft Excel 计算出表 6.2 中的结果。

表 6.2

轮 次	开价者	S 享有盈余	B 享有盈余
100	S	100	0
99	B	94	6
98	S	94.18	5.82
97	B	88.53	11.47
…	…	…	…
5	B	32.78	67.22
4	S	34.80	65.20
3	B	32.71	67.29
2	S	34.73	65.27
1	B	32.65	67.35

现在 B 的最佳初始开价应当是 232.65 元，从中他可以享有 67.35 元盈余，S 将接受这个开价而不愿不必要地推迟解决。注意盈余的分割近似地为 2∶1，对 B 有利。如果模型的其他假设都不改变，惟有轮次数从 100 增加到 150，这个对 B 有利的分割仍然基本如此，我们仍利用 Microsoft Excel 进行计算，并在表 6.3 中列出最后几个轮次的结果。

表 6.3

轮　次	开价者	S 享有盈余	B 享有盈余
6	S	34.10	65.90
5	B	32.05	67.95
4	S	34.09	65.91
3	B	32.04	67.96
2	S	34.08	65.92
1	B	32.04	67.96

表 6.3 显示的结果已经基本稳定。

§6.3 讨价还价的一般模型

两人讨价还价模型，如前所述，是由局中人交替地叫价以确定对盈余的如何分配。相当于两个人分一块蛋糕，在轮次 0,2,4,…，由局中人 B 提出一个分配方案 $(x,1-x)$，这里 x 可视作一块蛋糕的百分比，对此局中人 S 可以接受或拒绝。如果 S 接受则博弈结束，如果他拒绝 B 在 $2k(k=0,1,2,\cdots)$轮次的开价，那么在 $2k+1$ 轮次时他可以自己提出反建议方案，当然局中人 B 也可以接受或者反对，如果在某一轮次接受了 S 的开价，博弈宣告结束，否则双方交替地继续开出价格并继续就对方的方案表示自己的态度。理论上，可以作出开价的次数没有明文限制，我们称它为完全信息的

无限水平博弈。如果像§6.2那样，开价轮次数为固定的，这就是Ståhl于1972年研究的有限水平讨价还价博弈，它可以利用后退归纳思想求得解。将博弈从有限水平推广到无限水平的工作由Rubinstein于1982年进行。人们习惯地称本节所讨论的模型为Rubinstein-Ståbl讨价还价模型。通常总是考虑无耐心情况，我们所谈及的无耐心可以用折扣因子来表示。我们称某局中人再等一轮的花费是他下一轮分得盈余的某百分比 λ，这等价于这一轮的所得为下一轮所得的 $1-\lambda=\delta$ 倍。显然，δy（y 为他下一轮所得）相当于下一轮 y 在这一轮的"现时值"，δ 也就是我们以前提到过的折扣因子。在 R-S 模型中，总假设两个局中人具有非对称的无耐心，因此各表示为折扣因子 δ_B 与 δ_S，其中 $0<\delta_B<1$，$0<\delta_S<1$。

这个模型中存在大量Nasb均衡，例如策略剖面"局中人 B 总是要求 $x=1$ 且拒绝所有较小的分享；局中人 S 总是提供 $x<1$ 并接受任何开价"是一个Nash均衡。因为在这个策略剖面中，在任何一方策略给定的情况下，另一方倘想偏离不会给自己带来任何好处。但是，它不是子博弈完美均衡。理由很简单：问题出在折扣因子 δ_B 上。设想子博弈是从 S 拒绝 B 第一次开价 $x=1$ 开始，且 S 提供 B 以 $x>\delta_B$ 且 $x<1$，那么 B 应当接受。否则如果 B 拒绝 x，他至多在下一轮提出 $x=1$ 并被 S 所接受，然而那时的1只相当于现时的 δ_B，它显然小于局中人 S 提供给他的 x。与其下一次得到的1只相当于现时的 δ_B，不如现在就接受 $x>\delta_B$。

Rubinstein讨价还价模型有唯一的子博弈完美均衡，其结果的重要性足以让我们正式地以定理形式进行描述：

定理6.1　假设两个局中人 S 和 B，关于一块蛋糕（盈余）的分配用交替开价的办法进行讨价还价。局中人 B 作出第一次开价；可以作出的开价次数没有限制；两个局中人各自的折扣因子为 $0<\delta_B<1$ 与 $0<\delta_S<1$；当某局中人关于接受或拒绝某开价确实感

到无所谓时，则认为该局中人接受此开价。那么这个讨价还价博弈有唯一的子博弈完美均衡：局中人 B 立即提供给 S 以盈余的 $\delta_S \times (1-\delta_B)/(1-\delta_S \times \delta_B)$，而留给自己 $(1-\delta_S)/(1-\delta_S \times \delta_B)$，$S$ 接受此分配方案。

注：该策略剖面在由 B 叫价的子博弈部分，其叙述恰与原策略剖面的内容一样。如果子博弈是由 S 开价，那么相应的策略剖面叙述为：局中人 S 提供 B 以盈余的 $\delta_B \times (1-\delta_S)/(1-\delta_S \times \delta_B)$，留给自己 $(1-\delta_B)/(1-\delta_S \times \delta_B)$，$B$ 接受该分配方案。

按定理所述条件，只要局中人保持拒绝对方开价的状态，那么博弈在理论上继续到永远。因而这个博弈存在着无穷多个子博弈，正如我们所知，子博弈之一是原博弈，我们将它记作 G_1。假定 G_1 至少有一个子博弈完美均衡。在 G_1 所有子博弈完美均衡中，令 Q_B 表示局中人 B 最大盈利，q_B 表示 B 的最小盈利。完整博弈的第二个子博弈从 S 作出其第 1 次开价开始，不妨记作 G_2。记 Q_S 为 S 在 G_2 所有子博弈完美均衡中得到的最高盈利，相应地 q_S 为 S 的最小盈利。第 3 个子博弈——记作 G_3——开始于 B 对 S 作出第 2 次开价的阶段。不难想象，由于博弈是无限水平的，因此 G_3 可视作等价于 G_1，如果以 G_3 开始时刻的价值看作子博弈 G_3 的"现时价值"（当然将它折算成 G_1 开始时的现时价值的话，必须另外考虑折扣因子的作用），那么，显然 G_3 中所有子博弈完美均衡中 B 最高与最低盈利也等于 Q_B 与 q_B。

现在考虑在 G_1 中的第 1 个轮次中，B 应当如何开价。为了使 B 的第 1 次开价构成子博弈完美均衡中初始策略，那么这个开价应当有被 S 接受的机会，显然卖者接受的盈利至少应等于 $\delta_S \times q_S$。这是因为一旦博弈达到子博弈 G_2 时卖者可以保证具有盈利 q_S，而由于 q_S 的得到是在一个轮次后的"将来"，为了使这个盈利折算成现时价值，q_S 必须乘以折扣因子 δ_S。于是，相应地 B 至多可以得到

的是 $1-(\delta_S\times q_S)$。这相当于

$$Q_B\leqslant 1-(\delta_S\times q_S) \tag{6.1}$$

接下来我们想看看 q_B 有些什么约束条件。如果 B 提供给 S 任意大于或等于 $Q_S\times\delta_S$ 的盈余，S 肯定会接受它。因为 S 清楚地明白一旦到达子博弈 G_2，他至多得到 Q_S，而考虑到折扣因素，这个 Q_S 在 G_1 开始阶段的价值为 $\delta_S\times Q_S$，这蕴含着 B 自己至多可以得到 $1-(\delta_S\times Q_S)$，即

$$q_B\geqslant 1-(\delta_S\times Q_S) \tag{6.2}$$

我们可以将问题一开始就从 G_2 开始，利用与前面一样的讨论，不难得到

$$Q_S\leqslant 1-(\delta_B\times Q_B) \tag{6.3}$$

$$q_S\geqslant 1-(\delta_B\times Q_B) \tag{6.4}$$

(6.4)式的两端乘以负数($-\delta_S$)并再加上 1，则不等号反向，得

$$1-\delta_S\times q_S\leqslant 1-\delta_S(\delta_B\times Q_B)=1-\delta_S+\delta_S\times\delta_B\times Q_B \tag{6.5}$$

由(6.1)式、(6.5)式的左端大于等于 Q_B，于是我们有

$$Q_B\leqslant 1-\delta_S+\delta_S\times\delta_B\times Q_B \tag{6.6}$$

或者

$$Q_B\leqslant(1-\delta_S)/(1-\delta_S\times\delta_B) \tag{6.7}$$

类似地，在(6.3)式的两端乘以负数($-\delta_S$)并再加上 1，逆转其不等号，可得

$$1-(\delta_S\times Q_S)\geqslant 1-\delta_S+\delta_S\times\delta_B\times q_B \tag{6.8}$$

结合(6.2)式得

$$q_B\geqslant 1-\delta_S+\delta_S\times\delta_B\times q_B \tag{6.9}$$

或者

$$q_B\geqslant(1-\delta_S)/(1-\delta_S\times\delta_B) \tag{6.10}$$

比较(6.7)式与(6.10)式，可得

$$Q_B\leqslant(1-\delta_S)/(1-\delta_S\times\delta_B)\leqslant q_B \tag{6.11}$$

由定义，显然应有 $Q_B \geqslant q_B$，故(6.11)式中只可能全部成立等号，就是

$$Q_B = q_B = (1-\delta_S)/(1-\delta_S \times \delta_B) \qquad (6.12)$$

同样的逻辑推理不难证明

$$\begin{aligned} Q_S = q_S &= 1-(1-\delta_S)/(1-\delta_S \times \delta_B) \\ &= \delta_S(1-\delta_B)/(1-\delta_S \times \delta_B) \end{aligned} \qquad (6.13)$$

这样，我们实际上已经证明了定理的结论。回顾一下，博弈进行的无限水平在论证中占有关键作用，只有在无限水平的情况下，才能谈及 G_1 与 G_3 的等价性。在有限水平时，显然 G_1 与 G_3 通常不是等价的。

定理 6.1 的结论中，若固定 δ_S 而令 $\delta_B \to 1$，此时 $Q_B = q_B \to 1$。即 B 得到整块蛋糕。实际解释很容易：$\delta_B \to 1$ 即几乎等于 1，折扣因素越大，局中人再等待一个轮次的花费越小，若 $\delta_B = 1$，表明 B 在等待过程中没有任何花费，因此他显得“很有耐心”，耐心使他获得较多的盈余。同样，如果固定 δ_B：$0<\delta_B<1$，而令 $\delta_S \to 1$，蕴含 S 很有耐心，他将会得到整块蛋糕。还有一个有趣的情况：若 $0<\delta_B<1$，而令 $\delta_S=0$，此时局中人 B 也得到蛋糕。因为 $\delta_S=0$ 意味着局中人 S“极无耐心”，“缺乏远见”的局中人 S 将在今天接受任何正量的蛋糕分配而不会等到明天。反过来情况未必如此。若 $0<\delta_S<1$，而 $\delta_B=0$，此时局中人 S 得到 δ_S 而不是得到整块蛋糕。这是局中人 B 享有先动优势的原因，纵然 B 是一个缺乏远见、极无耐心的人，他也仍将获得 $1-\delta_S$ 的蛋糕。先动优势使得局中人 B 比局中人 S 处境好一些，体现这一点的另外一个例子是假使他们具有同样程度的耐心，$\delta_B=\delta_S=\delta \in (0,1)$，那么由定理 6.1，$B$ 应获得

$$Q_B = q_B = (1-\delta)/(1-\delta^2) = 1/(1+\delta) > 1/2 \qquad (6.14)$$

B 的蛋糕将大于 S 分得的蛋糕。

定理 6.1 指出，在无限水平讨价还价模型里，唯一的子博弈完

美均衡中 B 与 S 的所得，从公式来看，依赖于两个因素：

(1)两个局中人的折扣因子 δ_B 与 δ_S；

(2)在该模型中，谁首先行动。

我们已经分析了若干特殊的情况，得到的印象是一个有耐心的局中人可能获得较优厚的回报。在通常情况下常常有可能的确如此，但必须指出事情未必总是这样，例如，取 $\delta_B=0.70$，$\delta_S=0.75$，显然 S 比 B 有耐心，但是简单的计算表明买者 B 预期得到蛋糕的 53%，其原因在于 B 作为先动者，他仍享受着(小小的)先动优势。那么先动优势是否一定是优先于折扣因子(或耐心程度)的因素呢？其实也未必。如果我们所取每一轮次的时间间隔可以任意地短，先动优势将随之消失。为具体地讲清楚这一点，不妨令时间间隔长度为 Δ，置 $\delta_B=\exp(-r_B\Delta)$，$\delta_S=\exp(-r_S\Delta)$。设想 Δ 非常地接近于 0，近似地可以记 $\delta_B\approx1-r_B\Delta$ 与 $\delta_S\approx1-r_S\Delta$。此时，当 $\Delta\to0$ 我们有

$$\begin{aligned}(1-\delta_S)/(1-\delta_S\delta_B)&=r_S\Delta/[1-(1-r_S)(1-r_B\Delta)]\\&=r_S\Delta/[\Delta(r_S+r_B)-r_Sr_B\Delta^2]\\&=r_S/(r_S+r_B-r_Sr_B\Delta)\\&\to r_S/(r_S+r_B)\end{aligned}\qquad(6.15)$$

局中人的相对耐心程度决定了他们关于蛋糕的分享。特别地，如果两个局中人具有同等程度耐心，则有 $\delta_B=\delta_S$，从而 $r_S=r_B$，那么他们预期每人获得半块蛋糕。此时，先动优势荡然无存。

§6.4 序贯两人讨价还价的实验迹象

我们在前面详细介绍了两人讨价还价模型以及相应的一些结论。不过这样的模型常常是理想化了的。例如讨价还价的双方只是假定为理性人，因此，对于除讨价还价的顺序之外的其余一切因素，他们均处于平等且相同的地位。日常生活的经验告诉我们，其

实局中人的收入、地位、文化素质、性格等等都将直接地影响着两人讨价还价的最后结果。对于自由市场上同样质量的货，一般的工薪阶层，甚至一些因下岗而面临生活困难的人们，常常会与小贩进行较多轮次的讨价还价，而经济条件比较好的顾客相对来说讨价还价的轮次就要少一些。

已经进行大量实验以清晰地显示人们如何地在一些简单的场合讨价还价。实验的结果常常提出一些新的问题，它们涉及到刚才的一些讲法。

最简单的讨价还价例子是独裁者博弈(dictator games)。要求某当政者在他与另一个无名小卒之间分配一笔固定的钱。游戏规定，讨价还价只进行有限次，最后一次的指派分配人是当政者。根据 §6.1 的讨论，对于一个完全“自私自利”(即完全理性)的人来说，他将“几乎吃掉整块蛋糕”，因此理性的策略是所有钱归当政者自己——他享受“后动优势”。但是，大量实验结果表明，远远少于一半的当政者会选择这样的策略。那么，是否这足以证明当政者不是理性的呢？回答为“否”。这里我们必须注意到，博弈中局中人的盈利或效用起码地会依赖于其他一些因素——诸如名誉、地位或别的什么，此时收人只是考虑盈利时的一个小小因素。这一点与我们刚开始介绍博弈论时引进盈利概念所讲述的内容相符。博弈中盈利函数将直接影响结局，在“独裁者博弈”中建立的讨价还价模型考虑盈利纯粹是“钱”，对于无名平民来说，也许这样的盈利函数是符合实际的，对于当政者来说，“钱”或许不是他的全部盈利。

最后通牒博弈(ultimatum games)稍稍复杂一些。它要求第一个局中人提出对一笔固定的钱的分配。第2个局中人可以接受或者拒绝。如果拒绝分配方案，没有一个局中人得到任何东西。本博弈的一个特点是游戏规则允许第2个局中人具有一定的决断权，但提出方案的权利给了第1个局中人。毋庸置疑，理性的子博弈完美均衡是：第1个局中人提出自己取几乎整笔钱，留下一个正量 ε

(ε 可以是任意小的正数，但一旦选定后是固定的量)给第 2 个局中人。局中人 2 会接受这个分配方案，因为面对 ϵ 和 0 这两种可能结局的盈利，极大化自己的利益就等于接受第一个局中人的提议，这样至少可以使他得到 ϵ 正量的钱。然而，在此类实验中，通常“一半对一半”是第 1 个局中人的程式提案。而且，第 2 个局中人拒绝非零分配这种情况并不罕见。人们可以寻找各种理由来解释与理解这样的结局，例如第 2 个局中人以“宁可两个人什么也拿不到，自己也不愿只得到少量的 ϵ”来威胁第一个局中人。这个威胁未必“空头”，因为这对第 2 个局中人来说只损失 ε，而对第一个局中人来说则损失“几乎整块蛋糕”，第一个局中人不得不考虑这一威胁而以“五五开”使双方各有所得。当然，如果通过修改实验的结构也可能改变博弈的结局，譬如使局中人争取成为第一个局中人从而修改游戏的结构。

这两个讨价还价实验结果存在一些共同的问题，经常地局中人采取的策略使他们自己在财务方面更糟。例如最后通牒博弈中的第 2 个局中人，他竟然连正量的 ϵ 都不要，宁可什么也拿不到。用博弈论的术语来说，实验结果的多数不是子博弈完美均衡，局中人行为不是自私自利的。其中的一种解释是，博弈中的局中人认为极端的盈利差异是令人厌恶的，因此他们愿意放弃现利而拉平互相的盈利。Prasnikar 与 Roth 在 1992 年设计了一个实验以观察在称之为最佳机会博弈(the best shot game)中行为，看看在不同的模型中，局中人如何心甘情愿地避免盈利差距。在这个动态博弈中，先要求局中人 1 选择一个他愿意提供的整数量，表示为 Q_1。在观察到这个量之后，局中人 2 然后必须选择一个他自己愿意提供的整数量 Q_2。任何一个局中人提供的每一个单位将花费该局中人 0.82 个单位，返回给每一个局中人的收益由下式确定：

$$R_i(Q_1,Q_2)=\begin{cases}\max(Q_1,Q_2)-0.05\times\sum_{j=1}^{\max(Q_1,Q_2)-1} i & 如果\ \max(Q_1,Q_2)>1\\ 1 & 如果\ \max(Q_1,Q_2)=1\\ 0 & 如果\ \max(Q_1,Q_2)=0\end{cases}\tag{6.16}$$

该博弈的子博弈完美均衡是：局中人 1 提出整数量 0，局中人 2 提供整数 4。此时 $\max(Q_1,Q_2)=4>1$，因此局中人 1 的盈利为 $U_1(0,4)=R_1(0,4)-0\times0.82=4-0.05\times(1+2+3)=4-0.3=3.7$，而局中人 2 的盈利等于 $U_2(0,4)=R_2(0,4)-4\times0.82=4\times(1-0.82)-0.05\times(1+2+3)=4\times0.18-0.3=0.42$。我们用一步偏离准则来证明这个策略剖面是子博弈完美均衡。

在第 2 阶段中"局中人 2 提供 4"为固定不动的情况，我们试图证明局中人 1 不能偏离剖面，或者说，局中人 1 若提供 $Q_1\geq1$（Q_1 为整数），那么局中人 1 的盈利收获不会更好。设 Q_1 取作整数 1，2，3，4 中任何一个，那么 R_1 不会发生变化，但 $U_1=R_1-Q_1\times0.82=3.7-Q_1\times0.82<3.7$，局中人 1 的情况将会更糟。如果 $Q_1\geq5$，此时局中人 1 的盈利应为

$$\begin{aligned}U_1&=Q_1-0.05\times(1+2+\cdots+(Q_1-1))-0.82\times Q_1\\&=0.18Q_1-0.05\times Q_1(Q_1-1)/2\\&=Q_1(0.205-0.025Q_1)\end{aligned}\tag{6.17}$$

若假定 $Q_1>4$ 时连续变化，则易知 U_1 在 $Q_1=4.1$ 时达到极大值，当 $Q_1>4.1$ 时 U_1 呈递减状。特令 $Q_1=5$，此时 $U_1=0.4<3.7$，故 Q_1 取 5 及 5 以上整数时，U_1 更小，可见在第一阶段局中人 1 不能偏离上述策略剖面。现在考虑局中人 1 提供整数 0 时，局中人 2 是否可以不提供 4 而使自己更多获益。若 $Q_2=0$，由(6.16)，$U_2=0<0.42$。若 $Q_2=1$，则仍由(6.16)式，$R_2=1$，故 $U_2=1-0.82=0.18<0.42$。再考虑 Q_2 取大于 1 的整数，此时

$$\begin{aligned}U_2&=Q_2-0.05\times(1+2+\cdots+(Q_2-1))-0.82\times Q_2\\&=Q_2(0.205-0.025\times Q_2)\end{aligned}\tag{6.18}$$

此式与(6.17)式完全一致。视 Q_2 为大于 1 的数，易知 U_2 在 $Q_2=4.1$ 时达到极大值0.420 25，由于 Q_2 只能取整数，我们只要计算当 $Q_3=3$ 或 5 时，U_2 小于 0.42，就足以证明局中人 2 在第 2 阶段不应偏离。

$$U_2(0,3)=3\times(0.205-0.025\times3)=0.39<0.42$$

$$U_2(0,5)=5\times(0.205-0.025\times5)=0.4<0.42$$

在我们给出的子博弈完美均衡中，局中人 1 的盈利 3.7 与局中人 2 的盈利 0.4 之间存在着较大的差距。如果局中人 2 提供的整数小一些，那么他的花费相对也就少一些，有助于缩小两人盈利之间的差异。例如 Q_2 取 1，则局中人 1 得 1 而局中人 2 得 0.18，他们的盈利差就是局中人 2 提供 1 而付出的代价。更极端的一个情况是，局中人 2 也提供整数 0，于是两个局中人都两手空空，完全没有差异。这种情况有点像最后通牒博弈中那样，许多实验结果显示局中人 2 拒绝 ϵ 而使双方都一无所有。然而，实验结果却表明，在最佳机会博弈中的局中人的行为却紧密地与子博弈完美均衡预测的相一致。

第七章　宏观经济模型

众所周知，从长期来看，货币供应的增长决定了通货膨胀率，而且通货膨胀率也可影响产出增长与就业。假设国家中央银行关心通货膨胀率与就业水平，工人与雇主的行为明显地影响整个经济。而在中央银行、工人和雇主这三者之间的互相依存性正是本章试图通过建立博弈模型加以探索的。

工人与雇主倾向于签订长期合同，合同以工资的数字（称为票面数值）表示而不是以购买力表示（即实际价值）。这就要求工人在签订这样的工资合同时必须预测中央银行关于通货膨胀率的选择以及他们所能得到的工资的实际价值，若在完美信息与理性行为的模型中，已选择了子博弈完美均衡策略，其中工人将能够预测中央银行的行为，即使中央银行是在工人签订了他们的劳动合同之后再选择通货膨胀率的。

§7.1　宏观经济模型

工人与雇主试图预测的通货膨胀率并非完全如天气预报那样随机，它受到在中央银行控制之下的货币供应增长率的影响，因此工人与雇主的工资与就业决策将依赖于中央银行的货币增长决策。反转过来，银行的行为依赖于工人与雇主的工资与就业决策。这种相互依存性使得这三方面卷入一场博弈之中。

宏观经济建模强调经济的整体行为，忽略个体市场习性。因此

考虑如下因素：就业水平 L_t，价格水平 P_t，货币供应 M_t，货币工资 W_t，实际工资 R_t，预测的价格水平 P_t^e。每个变量的下标为 t，表示他们是随时间 t 而变化的。

我们将以自然对数来度量上述所有价格与其他量，它有如下若干好处：

(1)由于

$$\log(X_t)-\log(X_{t-1})\doteq x_t-x_{t-1}=\log(1+\frac{X_t-X_{t-1}}{X_{t-1}})$$

$$\approx\frac{X_t-X_{t-1}}{X_{t-1}} \tag{7.1}$$

这意味着在时间 t 与 $(t-1)$ 之间变量的自然对数差近似地等于在 $(t-1)$ 与 t 之间该变量的增长率。

(2)两个变量之比的对数相当于它们的对数之差：

$$\log(\frac{X}{Y})=\log X-\log Y \tag{7.2}$$

(3)1 的自然对数等于零。

现在我们定义下述变量

$$\pi_t=p_t-p_{t-1}=\log P_t-\log P_{t-1}\approx\frac{P_t-P_{t-1}}{P_{t-1}} \tag{7.3}$$

π_t 为价格通货膨胀率；

$$\pi_t^e=p_t^e-p_{t-1}=\log P_t^e-\log P_{t-1}\approx\frac{P_t^e-P_{t-1}}{P_{t-1}} \tag{7.4}$$

π_t^e 为预期的价格通货膨胀率；

$$g_t=m_t-m_{t-1}=\log M_t-\log M_{t-1}\approx\frac{M_t-M_{t-1}}{M_{t-1}} \tag{7.5}$$

g_t 为货币供应的增长率。

如下定义的 r_t 作为对数形式的实际工资：

$$r_t=\log R_t=\log\frac{W_t}{P_t}=\log W_t-\log P_t=w_t-p_t \tag{7.6}$$

当 $t=1$ 时，某些量标准化：$M_1=1$ 与 $P_1=1$，故

$$m_1=0, p_1=1 \tag{7.7}$$

于是，

$$\pi_2=p_2-p_1=p_2 \tag{7.8}$$

$$\pi_2^e=p_2^e-p_1=p_2^e \tag{7.9}$$

$$r_2=w_2-p_2=w_2-\pi_2 \tag{7.10}$$

注意到我们对宏观经济的建模由于与时间 t 有关，显然是动态博弈。我们的建模将博弈进程假设在两个时间期间进行。在第一期间，工人选取他们将在第二期间工作的货币工资。他们是在不知道下一期间的价格水平情况下作出决策的。在第二期间的开始时，中央银行选择通货膨胀率。然后雇主选取就业水平，根据以上所述过程，不难描述博弈树如图 7.1 所示。

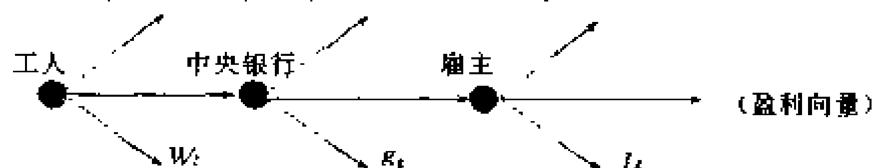


图7.1

图 7.1 中，工人选择的 W_t、中央银行选择的货币增长率 g_t(从而反映了通货膨胀)以及雇主选择的就业水平 L_t 等，其选择范围常常具有或几乎具有连续统势，因此博弈树中无法一一表示，只能用虚线箭头形式代之。

中央银行是通过控制货币供应增长来达到控制通货膨胀率的，尽管经济学家尚未建立货币增长与通货膨胀之间的因果关系，我们将采取极端的货币专家立场并假设所有的价格通货膨胀是因货币增长引起的。即

$$\pi_2=\Phi(g_2) \tag{7.11}$$

其中，$\Phi(\cdot)$是严增函数。

来自世界各国的一些历史数据非常有利于这两者之间存在着因果关系的假设，因此我们今后假定中央银行直接地选择通货膨胀率 π_2，这个有一定证据的合理假设有助于我们建立博弈模型。

从图 7.1 可以看出，博弈存在着三个局中人：工人、雇主与中央银行。工人的策略为货币工资 w_2，这里使用记号 w_2 是因为该工资是在博弈第二个期间内的工作所得。中央银行现在选择的策略由通货膨胀形式 $\pi_2(w_2)$ 组成，而不再用 g_2。通货膨胀率 π_2 依赖于工人所选择的工资 w_2 这一事实是本博弈模型里中央银行取糟糕行为的重要原因。至于雇主的策略，则由行为 $l_2(w_2,\pi_2)$ 的雇用规则组成。

还剩下一个重要的待定因素就是博弈树末端各方的盈利。我们将分别讨论：

工人作为一个整体，其效用极大化将从两个方面加以考虑，第一是收入，第二是空闲程度（没有事干）。很清楚，如果实际工资固定，那么增加雇佣必定导致工人作为整体来说增加了实际收入，相应地减少了失业。因此工人的盈利或效用（这里似乎用"效用"两字更妥）与提供的劳力水平有关，而该因素显然与实际工资 r_t 有关。我们记 $l^s(r_t)$ 为在给定实际工资下，劳力供给的极大化水平，即劳力供给曲线。为简单起见，假设

$$l^s(r_t)=\alpha\cdot r_t \tag{7.12}$$

其中 α 是一个常数，它相当于实际工资的 1%变化所导致的劳力供给的百分数变化。已经知道，由雇主最终选择的雇佣水平为 l_2，最大的劳力供应与 l_2 相差无几是工人的愿望。因为这预示着工人的空闲或失业可能减少而实际收入有所提高。因此工人的效用函数可以定义作

$$\begin{aligned}u_w(w_2,l_2,\pi_2)&=-(l_2-\alpha r_2)^2\\&=-(l_2-\alpha(w_2-\pi_2))^2\end{aligned} \tag{7.13}$$

中央银行关心就业水平与通货膨胀这两个因素，对此它有自己的目标值，记作 $\tilde{l}_2$ 与 $\tilde{\pi}_2$，并且试图利用它的政策手段——π_2——指导 l_2 与 π_2 趋于它们的目标值。正由于此，假定中央银行的效用函数为

$$u_c = -(l_2 - \bar{l}_2)^2 - \mu(\pi_2 - \tilde{\pi}_2)^2 \tag{7.14}$$

外生常数 μ 告诉人们，这两个因素与目标值的接近程度对中央银行不是处于同等位置，它度量了中央银行认定就业目标对于通货膨胀目标的相对重要性。

雇主的策略是选取 l_2，这与使他利润达到极大化时的劳力需求并不一定相符，劳力需求函数当然与实际工资 r_t 有密切关系，由于在本博弈模型中所有因素都用自然对数来度量，因此劳力需求函数也不例外地使用自然对数，记作 $l^d(r_t)$，假定

$$l^d(r_t) = -\eta \cdot r_t \tag{7.15}$$

其中 η 又是一个外生常数。雇主希望极大化自己的效用，他自然希望所选择的 l_2 与 $l^d(r_t)$越接近越好，根据这个观点，雇主的效用函数可以定义作

$$\begin{aligned} u_E &= -(l_2 - l^d(r_t))^2 \\ &= -(l_2 - \eta(\pi_2 - w_2))^2 \end{aligned} \tag{7.16}$$

当劳力供给量等于劳力需求量时，无疑劳力市场处于均衡状态。由 $l^s(r_t)$与 $l^d(r_t)$的定义可知，只有 $r_t = 0$ 才可能发生此事。在该均衡实际工资的雇佣水平称为雇佣的自然水平(natural level of employment)。我们仅需适当地选取变量单位，不难使得雇佣的自然水平等于 1，其对数为 0。当劳力市场处于均衡时，也许人们会预期此时没有一个人失业。其实不然，这种预期是不正确的。即使当劳力市场处于均衡状态，一些工人将继续寻找那些高于他们迄今为止所得收入的工作。这些工人是自愿失业的，他们的失业称作摩擦失业(frictional unemployment)。至于现代竞争性经济是否可能存在一个雇佣低于自然水平的宏观均衡是一个重要的尚待探索的问题。在那种均衡状态，一些工人将愿意干付给他们当前实际工资的任何工作，然而雇主不雇用他们。当然，这些工人是不自愿地失业了。

§7.2 宏观动态博弈的均衡解

注意到我们建立的宏观经济动态博弈具有完美信息，因此可以利用后退归纳法确定子博弈完美均衡。

从最后一个阶段着眼考虑，行动的局中人是雇主。假设工人选择了货币工资 w_2、中央银行选择了通货膨胀率 π_2，那么在给定 w_2 与 π_2 条件下，雇主想极大化自己的效用 u_E，唯一可选择的策略为 l_2，且由(7.16)式易知，最优就业水平 $l_2^*(w_2,\pi_2)$ 应当等于 $\eta(\pi_2-w_2)$。这就是经济学的总就业函数(aggregate employment function)：

$$l_2^*(w_2,\pi_2)=\eta\cdot(\pi_2-w_2) \tag{7.17}$$

后退到中央银行行动的阶段。也就是说，中央银行需要极大化 u_c，u_c 依赖于 w_2、l_2 与 π_2，其中 π_2 是待定的，而 $l_2=l_2^*(w_2,\pi_2)$。因此，中央银行应当如此地选择 π_2，使得下式极小化：

$$[\eta(\pi_2-w_2)-\tilde{l}_2]^2+\mu(\pi_2-\tilde{\pi}_2)^2 \tag{7.18}$$

在(7.18)式中对 π_2 求导数并使之为零：

$$\eta[\eta(\pi_2-w_2)-\tilde{l}_2]+\mu(\pi_2-\tilde{\pi}_2)=0 \tag{7.19}$$

解(7.19)式得

$$\pi_2^*(w_2)=\frac{\mu\tilde{\pi}_2+\eta\tilde{l}_2+\eta^2w_2}{\mu+\eta^2} \tag{7.20}$$

最后我们后退到工人行动的阶段：工人对于工资合同的选择依赖于他们关于就业和通货膨胀率的预测，在宏观经济文献中，对通货膨胀的预测——π_2^e——称为预期通货膨胀。工人怎样去预测通货膨胀呢？理性支配他们通过向前看并预见由中央银行的最优通货膨胀政策将产生的通货膨胀率以形成他们的预期。直截了当一些，就是

$$\pi_2^e=\pi_2^*(w_2) \tag{7.21}$$

假如工人相信这个办法，则称为具有理性的预期通货膨胀(ratio-

nal inflationary expectations)。芝加哥大学的 Robert Lucas 教授由于使用理性预期的先驱工作而获得 1995 年诺贝尔经济学奖。将 $\pi_2^*(w_2)$ 与 $l_2^*(w_2,\pi_2)$ 这两个解(易见,它们是除工人行动之后的子博弈中的完美均衡解)代入工人的效用函数:

$$
\begin{aligned}
u_w &= -[l_2^*(w_2,\pi_2^*(w_2)) - \alpha(w_2 - \pi_2^*(w_2))^2 \\
&= -[\eta(\pi_2^*(w_2) - w_2) - \alpha(w_2 - \pi_2^*(w_2))^2 \\
&= -(\eta+\alpha)^2[\pi_2^*(w_2) - w_2]^2
\end{aligned}
\tag{7.22}
$$

显然,当 $\pi_2^*(w_2^*) = w_2^*$ 时 u_w 达到极大,此时有(7.20)式得

$$w_2^* = \tilde{\pi}_2 + \frac{\eta}{\mu}\bar{l}_2 \tag{7.23}$$

$$\pi_2^*(w_2^*) = w_2^* = \tilde{\pi}_2 + \frac{\eta}{\mu}\bar{l}_2 \tag{7.24}$$

于是我们得到宏观经济动态博弈的子博弈完美均衡解为

$$\left\{w_2^* = \pi_2^* = \tilde{\pi}_2 + \frac{\eta}{\mu}\bar{l}_2, l_2^* = \eta(\pi_2^* - w_2^*) = 0\right\} \tag{7.25}$$

由(7.10)式,在第二期间的均衡实际工资(取自然对数) r_2^* 等于零。因此劳力供给与劳力需求公式给出

$$l_2^d(r_2^*) = l_2^s(r_2^*) = 0 \tag{7.26}$$

这表示劳力市场处于均衡状态。

当且仅当就业目标 $\bar{l}_2 = 0$ 时,中央银行同时也达到了通货膨胀的目标 $\tilde{\pi}_2$,这从(7.24)式立即可以得到。但是,存在着许多原因使得中央银行认为雇用的自然水平从全社会效益来说并不是最优的,因此他们将目标 $\bar{l}_2$ 定在自然水平之上。于是 $\bar{l}_2$ 严格地说是一个正数。由于 $\bar{l}_2$ 不是自然水平,通货膨胀率 π_2^* 也就高于通货膨胀目标值 $\tilde{\pi}_2$。在这个动态博弈中,低通货膨胀货币政策不是时间一致的(time consistent):当中央银行到了选择通货膨胀率时,它将发现遵循原来设想的低通货膨胀率 $\tilde{\pi}_2$ 不是最佳的,代之以最高的通货膨胀率 π_2^*,它是子博弈完美均衡解,当然是最佳的。

第八章　重复博弈

最简单的也是最特殊的可观察行动的多阶段博弈是重复博弈，在§5.3中我们已经初步涉及。假设在海滩占位博弈中的两个小贩，每天选择一个矿泉水价格进行价格竞争，日复一日地，他们进行着重复的价格博弈。类似的例子在经济学中不胜枚举。但是，诸如投资、学习等重要经济活动却不能用重复博弈建模，原因在于前一轮的投资行为或学习将影响到目前阶段中的可行行为空间与盈利(或效用)函数的形状与性质，这样，这一轮(或称周期)的博弈就不是前一轮(或周期)博弈的重复。所谓重复博弈就是在每一周期中局中人面对相同的博弈——经典的重复博弈理论中相同的博弈是指相同的局中人集合、相同的可行行为空间或策略空间，以及相同的盈利或效用函数，我们称它为每周期中的阶段博弈(stage game)。如果每一周期的末期，局中人的行为可以被观察到，那么局中人就有可能在对手过去的行为基础上，在下一周期采取自己(有条件的)策略。我们将会看到，这一特点可能导致的均衡结局会出现一次博弈中的非均衡结局。

我们先从有限重复博弈开始，逐步展开讨论。

§8.1　有限重复博弈

顾名思义，有限重复博弈就是阶段博弈重复实施有限(T)次。先不妨令 $T=2$，考虑如图8.1所示的完全信息静态博弈。

		局中人 2	
		L_2	R_2
局中人 1	L_1	1,1	5,0
	R_1	0,5	4,4

图 8.1

显然，图 8.1 属于囚徒窘境型博弈，它只有一个 Nash 均衡解(L_1, L_2)，有效结局(R_1, R_2)不是均衡解。假设该博弈实施两次，两阶段重复博弈中每个局中人的盈利相当于各个阶段盈利之和（或者平均数），考虑到第二阶段可能存在的折扣因子 δ，因此应当为第一阶段的盈利加上 δ 倍的第二阶段盈利（若两阶段之间间隔时间很短，可认为 $\delta=1$）。为求子博弈完美均衡解，再一次借助于后退归纳法。显然，第二阶段的唯一 Nash 均衡是(L_1, L_2)，盈利向量是(1,1)。因此，如果子博弈完美均衡存在的话，其第二阶段结局必定是(L_1, L_2)，所得盈利的现时值为(δ, δ)。不管子博弈完美中第一阶段取何种结局，该结局的盈利向量加上(δ, δ)就得到局中人的子博弈完美均衡盈利。根据此分析，我们可以在图 8.1 的盈利矩阵的各个结局中都加上(δ, δ)，得到如图 8.2 所示的新盈利矩阵。

		局中人 2	
		L_2	R_2
局中人 1	L_1	$1-\delta, 1+\delta$	$5+\delta, \delta$
	R_1	$\delta, 5+\delta$	$4+\delta, 4+\delta$

图 8.2

图 8.2 中的博弈也有四个结局，而每个结局(x, y)对应了两阶段重复博弈的行动系列中的一个行动：$\{(x, y), (L_1, L_2)\}$，相应于图 8.2 中(x, y)的盈利向量就是局中人在该行动中的所得。累次取优法告诉我们，(L_1, L_2)是唯一 Nash 均衡，因此我们得到了唯一的

子博弈完美 Nash 均衡：$\{(L_1,L_2),(L_1,L_2)\}$。

如果把图 8.1 中 (L_1,L_2) 的盈利改为 $(2,1)$，则我们在盈利矩阵的各个结局中加上$(2\delta,\delta)$。以上分析的结论依然成立。这是因为，无论是累次取优，还是使用划线法求 Nash 均衡解，都是在给定其他局中人所取策略的条件下，比较同一局中人采取不同策略时的盈利大小。因此，在每一个局中人的各种可能盈利上加一个常数（不同局中人的盈利所加常数可以不相同）之后，博弈的 Nash 均衡仍为"新"博弈的均衡结局。

两阶段重复博弈的上述结论很容易推广到任意有限(T)次重复博弈，我们实际上处理这样的"一次性"完全信息静态博弈，其盈利矩阵相当于被重复的博弈的盈利矩阵中各元素扩大到$(1+\delta+\delta^2+\cdots+\delta^T)$倍。新的"一次"博弈的 Nash 均衡仍为$(L_1,L_2)$（继续以图 8.1 为例），这蕴含着 T 次重复博弈有着唯一的子博弈完美 Nash 均衡：$\{(L_1,L_2),\cdots,(L_1,L_2)\}$。

倘若再将两人博弈的情况稍微扩大到 n 人博弈，我们可以正式地叙述上述结论如下：

定义 8.1 令 $G=\{A_1,\cdots,A_n;u_1,\cdots,u_n\}$表示 n 个局中人的完全信息博弈，对 G 重复若干次，称 G 为阶段博弈。给定阶段博弈 G，令 $G(T)$表示 G 实施 T（T 为大于 1 的整数）次的重复博弈。在某次阶段博弈开始之前，所有已采取过的前面阶段的行动都可以观察到。局中人在 $G(T)$的盈利或效用简单地为来自 T 个阶段博弈盈利现时值之和。

注：$G(T)$的盈利也可定义为 T 个阶段博弈盈利现时值的平均，它与现时值之和仅相差常数因子 $1/(1+\delta+\delta^2+\cdots+\delta^T)$，这并不影响重复博弈的子博弈完美结局。

定理 8.1 如果阶段博弈 G 有唯一的 Nash 均衡，那么对任意有限次 T，重复博弈 $G(T)$有唯一的子博弈完美结局：在每一阶段

取 G 的 Nash 均衡策略。

定理 8.1 的证明就是前面的一段论述，利用后退归纳法并将每个阶段的唯一 Nash 结局盈利"糅合"进第一阶段博弈的盈利矩阵，得到一个新的"一次性博弈"，其唯一的 Nash 均衡就是重复博弈的子博弈完美均衡解。其实，通俗地解释定理 8.1.在有限重复博弈中，每一次博弈，摆在局中人面前只有"华山一条路"，偏离此路就会迷失方向并招致损失。于是也就锁定了每个局中人必然采取的策略。根据定理 8.1，如果囚徒窘境重复 T 次，那么"总是坦白"将是每个局中人采取的子博弈完美策略。

现在，进一步阐述定理 8.1：

(1)我们以图 8.1 为例引进有限重复博弈的定义与定理 8.1，注意到图 8.1 仅存在唯一纯策略 Nash 均衡。而定理 8.1 中要求的唯一 Nash 均衡可以是混合策略均衡。我们借助于猜谜博弈(见图 8.3)来讲述这个问题。

		局中人 2：1	局中人 2：2
局中人 1	1	1,1	−1,1
	2	−1,1	1,−1

图 8.3

正如我们已经指出的，猜谜博弈不存在纯策略解，它具有唯一的混合策略 Nash 均衡{(1/2,1/2),(1/2,1/2)}，相应的期望盈利为(0,0)。不难知道，最后的一次博弈仍如图 8.3，于是我们可以得到 T 次重复猜谜博弈的子博弈完美均衡：每个局中人总是以 1/2 概率伸出 1 根手指，以 1/2 概率伸出 2 根手指。实验证明，这的确是猜谜游戏中双方采取的最佳策略。

(2)定理 8.1 表述中，我们只规定阶段博弈 G 是完全信息博弈，没有硬性指定 G 一定是静态的。其实，假如 G 是完全且完美信息动态博弈时，定理 8.1 仍成立类似的结论。设 G 具有唯一的"后

退归纳"结局，那么 $G(T)$ 有唯一的子博弈完美 Nash 均衡：每一周期(注：这里使用术语"周期"，原因在于 G 是动态的，比如它可以是可观察行动的多阶段博弈)都取 G 的"后退归纳"结局。理由很简单，在每个周期，只有这条"路"是合理的预测结局。

(3)要求阶段博弈有唯一 Nash 均衡，是定理 8.1 中除"有限"之外另一个关键性条件。设想阶段博弈具有多重 Nash 均衡，那么人们有理由要问，利用后退归纳法从最后一个阶段博弈着手考虑问题，哪一个 Nash 均衡作为"后退"的出发点比较合理呢？显然，硬性规定一个 Nash 均衡是毫无道理的，况且后退到倒数第二个阶段博弈时，又遇到了同样的问题，是取与适才相同的均衡还是取另一个均衡，这些都是在阶段博弈只有唯一 Nash 均衡时不可能出现的问题。可以想象，如果阶段数目越多和 Nash 均衡的个数越多，似乎"后退之路"也就越多，可惜这并不构成"条条大道通罗马"的格局，相反，使我们预测 $G(T)$ 的合理结局时遇到了棘手的麻烦。§5.3 中我们已经简单地讨论过这方面的例子。这里我们仍以两阶段博弈为例对此进行探讨。

对图 8.1 中的博弈人为地给两个局中人各添加一个纯策略，其盈利矩阵如图 8.4 所示。

局中人 2

局中人 1		L_2	R_2	Q_2
	L_1	1,1	5,0	0,0
	R_1	0,5	4,4	0,0
	Q_1	0,0	0,0	3,3

图 8.4

实际上，我们人为地制造了一个纯策略 Nash 均衡 (Q_1,Q_2)。由于只讨论两阶段重复博弈，为方便起见，令折扣因子 $\delta=1$。对于第一阶段博弈中的每一个可能结局 (x,y)，根据子博弈完美均衡

的定义，局中人必定预期第二阶段博弈出现 Nash 均衡结局。问题在于我们有不止一个 Nash 均衡，即使仅限于纯策略 Nash 均衡，也有两个：(L_1, L_2)与(Q_1, Q_2)。现在，局中人可能预期在不同的第一阶段结局后面将跟着第二阶段不同的均衡结局。两阶段重复博弈的子博弈完美均衡与局中人的预期将会有密切的关系。

例如，假设局中人预期第一阶段的结局为(R_1, R_2)的话，则第二阶段结局会是(Q_1, Q_2)；而如果第一阶段出现除(R_1, R_2)之外其余 8 个结局之一，那么在第二阶段预期结局是(L_1, L_2)。这种预期可以看作是局中人谈判的结果，如果第一阶段出现(R_1, R_2)，则第二阶段的(Q_1, Q_2)是对双方的奖励。若第一阶段出现了非(R_1, R_2)，这表示有人违反谈判协议，于是在第二阶段得到惩罚。基于局中人的上述预期（或者说基于上述“条件”），两阶段重复博弈的可能策略剖面为

$$\{(R_1, R_2), (Q_1, Q_2)\} \text{和} \{(x, y), (L_1, L_2)\}$$

其中 $$(x, y) \neq (R_1, R_2)$$

这种情况下，又可以将两阶段“粿合”成一次博弈，其新的盈利矩阵如图 8.5 所示。

局中人 2

局中人 1		L_2	R_2	Q_2
	L_1	2,2	6,1	1,1
	R_1	1,6	7,7	1,1
	Q_1	1,1	1,1	4,4

图 8.5

图 8.5 实际上是将(3,3)加到图 8.4 中(R_1, R_2)对应的格子，而其余 8 个格子均加上(L_1, L_2)的盈利(1,1)。新的一次博弈有三个纯策略 Nash 均衡：(L_1, L_2)，(Q_1, Q_2)与(R_1, R_2)。这三个 Nash 均衡分别对应于原重复博弈的子博弈完美均衡。(L_1, L_2)对应于

$\{(L_1,L_2),(L_1,L_2)\}$，(Q_1,Q_2)对应于$\{(Q_1,Q_2),(L_1,L_2)\}$，这些策略剖面中的各个阶段结局都是阶段博弈的 Nash 均衡。余下的(R_1,R_2)对应于两阶段重复博弈的子博弈完美结局$\{(R_1,R_2),(Q_1,Q_2)\}$。产生的一个与前两者不同的结果是，在第一阶段可以达到原阶段博弈中有效的非 Nash 均衡(R_1,R_2)。现对更一般问题进行叙述来总结这个"有趣"的结果：

如果阶段博弈 $G=\{A_1,\cdots,A_n;u_1,\cdots,u_n\}$是具有多重 Nash 均衡的完全信息静态博弈，那么可能（但不必）存在重复博弈 $G(T)$的子博弈完美均衡结局，其中对任意 $t>T$，在 t 阶段的结局并不是 G 的 Nash 均衡。

上述例子中的子博弈完美均衡是根据"局中人基于第一阶段结局如何预期第二阶段结局"而获得的。如前所述，这是双方"谈判"达成的协议。这里面蕴含着如下有意义的事实：在重复博弈中，对于将来行为的可信承诺（或威胁）可以影响现时的行为。例如，如果局中人采取策略产生第一阶段的有效非 Nash 均衡(R_1,R_2)，那么承诺在第二阶段以(Q_1,Q_2)给予奖励（因为(Q_1,Q_2)的盈利明显地大于(L_1,L_2)的盈利），正是这种承诺的可信性，局中人预期(R_1,R_2)后面将会是(Q_1,Q_2)。读者也许会问，既然是承诺奖励，那么为什么不讲好以(R_1,R_2)的盈利（4，4）而却用(Q_1,Q_2)的（3，3）呢？理由相当简单：(R_1,R_2)不是阶段博弈的 Nash 均衡。由于第二阶段是最后一个阶段，不存在下一阶段进行"惩罚"的可能，倘若"讲好"取(R_1,R_2)的话，局中人为了极大化自己的盈利，有可能采取欺诈行为从而偏离(R_1,R_2)。显然局中人 1 或 2 都希望偏离 R 而取 L，以使自己的盈利从 4 增加到 5，但其最后结局却成为(L_1,L_2)，虽然它也是 Nash 均衡，当然比另一个 Nash 均衡(Q_1,Q_2)在盈利方面差得多。

基于这种讲法，读者会认为在第一阶段取(R_1,R_2)之后预期第二阶段为(Q_1,Q_2)是合理的。但紧跟着想不通的问题是，为什么

局中人在其他 8 个结局之后会预期 (L_1,L_2) 呢? 在第二阶段取 (L_1,L_2),它的盈利仅为(1,1),虽然它是 Nash 均衡,但看起来似乎只有蠢人才会做这样的事,因为另外的 Nash 均衡 (Q_1,Q_2)从盈利角度是个更好的选择。因此人们可以这样地进行推理:不管第一阶段出现什么样的结局,过去的事情就让它过去吧,在重新开始的第二阶段博弈中预期结局应取 Nash 均衡(Q_1,Q_2),毫无疑义它是受人欢迎的。以上说法似乎很有理,然而按照这种逻辑,局中人预期第一阶段的任何结局后将跟着第二阶段 Nash 均衡结局(Q_1, Q_2)。在这种预期下,反过来预测或激励在第一阶段采取(R_1,R_2)也许是灾难性的。因为第二阶段千篇一律地取(Q_1,Q_2),使得第二阶段结果对第一阶段没有任何威胁或承诺(奖励)。于是在第一阶段局中人为了极大化自己的盈利会主动地偏离(R_1,R_2)。显然 L_i 是局中人 i 关于对手取 R_j 时的最佳反应。预测的结局将不是(R_1, R_2),而是(L_1,L_2)。

也有些读者会提出,上述例子中阶段博弈的不同 Nash 均衡具有不同的盈利,从而产生了承诺或威胁之类的问题。其实,即使 Nash 均衡的盈利或效用完全一样,也会存在一些在阶段博弈只存在唯一 Nash 均衡时所没有的问题。为说明这个问题,我们人为地构造如图 8.6 所示的阶段博弈。

		局中人 2		
		L_2	R_2	Q_2
	L_1	3,3	0,1	0,0
局中人 1	R_1	1,0	3,3	1,1
	Q_1	0,0	1,1	3,3

图 8.6

G 显然存在三个纯策略 Nash 均衡 (L_1,L_2),(M_1,M_2) 与 (R_1,R_2),它们具有相同的盈利向量(3,3)。局中人无论预测在第二阶段出现哪一个 Nash 均衡,在将第二阶段预期盈利加到第一

阶段博弈的盈利矩阵时，相当于图 8.6 的每一个格子均加上(3，3)。无疑，新的一次博弈的 Nash 均衡仍为(L_1,L_2)，(M_1,M_2)和(R_1,R_2)，它们所蕴含的子博弈完美结局却是多种多样的。但是所有可能的子博弈完美具有相同的盈利(6，6)。那么如何预期第二阶段会是哪一个均衡结局呢？看起来，协商或重新协商是一种可能的解决办法。

阶段博弈 G 具有多重 Nash 均衡所引出的重复博弈中的问题，在无限重复博弈中也会遇到，在本章中我们将另立一节专门讨论 Pareto 完美与重新协商(或谈判)等处理此类问题的一些方法。

§8.2 无限重复博弈与无名氏定理

如果囚徒窘境实施一次或者有限次，我们已经知道，两个囚徒“总是坦白”构成了子博弈完美均衡。倘若该博弈不断地重复实施，每次博弈前可以看到以前各次曾采取的行动，我们就可以认为这是无限重复的博弈。本节的讨论将告诉人们一个有趣但吃惊的结果：两个囚徒“总是抗拒”是无限重复博弈的子博弈完美均衡，尽管{抗拒，抗拒}是阶段博弈的有效却非均衡结局。我们已经介绍过，在有限重复博弈中，如果 G 具有多重 Nash 均衡，可能存在这样的子博弈完美，对任意 $t<T$，t 阶段的结局不是 G 的 Nash 均衡。而无限重复博弈的这个有趣结果比起有限重复博弈来，似乎有过之而无不及。纵然阶段博弈 G 有唯一 Nash 均衡，无限重复博弈可能存在这样的子博弈完美均衡，其中没有一个阶段结局是 G 的 Nash 均衡。这一结果更能体现出，在无限重复博弈中，关于未来行动的可信威胁与承诺可能影响当前的行为。

本节将重点研究无限重复博弈。由于博弈无数次地重复，从而不存在最后一次博弈，后退归纳法将无法在这里施展身手。因此我们先从建模着手研究。

1．无限重复博弈模型

模型中的基本点是阶段博弈G。假设G具有有限个局中人$i=1,2,\cdots,n$，他们同时行动。每个局中人i具有有限行动（或策略）空间A_i，令A为n个A_i的乘积空间：

$$A=X_{i=1}^{n}A_i=A_1\times A_2\times\cdots\times A_n$$

局中人i在阶段博弈G中的盈利或效用函数g_i实质上是从A到实数空间的函数。令$\mathscr{A}_i$为在A_i上的概率分布空间。

我们仍然考虑可观察行动的重复博弈，即局中人在每一局期结束时观察到已实施的行动。令$a^t=(a_1^t,\cdots,a_n^t)$为t周期中的行动。那么$h^t=(a^1,a^2,\cdots,a^{t-1})$就是$t$周期博弈开始时的历史。

定义重复博弈，必须确定局中人的策略空间和盈利函数。令s_i^t表示局中人i在t周期博弈时基于历史h^t所采取的行动，那么重复博弈中局中人i的纯策略s_i为序列，即

$$s_i=(s_i^1,s_i^2,\cdots,s_i^t,\cdots) \tag{8.1}$$

而混合（行为）策略σ_i为一系列σ_i^t，其中σ_i^t表示局中人i基于历史采取混合策略$\alpha_i\in\mathscr{A}_i$。

现在来确定局中人的盈利函数，假定折扣因子为δ：$0<\delta<1$，故可记相应的无限重复博弈为$G(\delta,\infty)$，或简单地记作$G(\delta)$。若记$\sigma^t=(\sigma_1^t,\sigma_2^t,\cdots,\sigma_n^t)$，局中人$i$的目标函数是极大化正则和

$$u_i=E_\sigma(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}g_i(\sigma^t(h^t)) \tag{8.2}$$

其中运算子E_σ是指关于由策略剖面σ所产生的无限历史上的分布求期望。因子$(1-\delta)$用来以相同的单位测度阶段博弈与重复博弈盈利：每个周期盈利均为1个单位的正则（或标准化）值为1。

因为我们需要研究子博弈完美均衡，因而必然涉及到真子博弈的盈利。注意到每个周期实际上是一个真子博弈的开始，对于任何策略剖面σ和历史h^t，我们可以计算从t周期开始的局中人的期望盈利——“持续盈利”，并且再以t周期时刻的单位测度对从t

开始的持续盈利进行标准化。于是，从 t 开始的持续盈利为

$$(1-\delta)\sum_{\tau=t}^{\infty}\sigma^{\tau-t}g_i(\sigma^\tau(h^\tau)) \tag{8.3}$$

显然，这样的再标准化使得从 t 周期开始，每个周期接受 1 单元盈利的局中人的持续盈利为 1。虽然在无限重复博弈中我们常考虑 $0<\delta<1$，但是有时候也考虑 $\delta=1$ 的情况，我们从讨价还价模型这一章知道，此时称局中人"完全地耐心"。在对完全耐心情况建模时，对盈利函数常采用平均准则，其中每个局中人 i 的目的是极大化

$$\lim\inf_{T\to\infty}E(1/T)\sum_{t=1}^{T}g_i(\sigma^t(h^t)) \tag{8.4}$$

这种平均准则蕴含着局中人不但对盈利的时间性不关心，而且对任意有限个周期的盈利也不关心。

已知有限的完全信息静态博弈若重复无数次，且在每周期博弈开始前都可以观察到以前周期的所有已采取过的行动，在确定了上述有关局中人的策略空间与盈利函数之后，实际上我们在前面已经建立了可观察行动的无限重复博弈模型。有一个显然的观察结果，尽管简单但是十分有用，我们用结论的形式罗列如下：

观察结果 8.1　设 $\alpha^*=(\alpha_1^*,\cdots,\alpha_n^*)$ 是阶段博弈的 Nash 均衡，那么策略"每个局中人 i 从现在开始采取策略 α_i^*"是子博弈完美均衡。再者，如果阶段博弈 G 有 m 个静态 Nash 均衡 $\{\alpha^j\}_{j=1}^m$，那么令 $j(t)$ 为时间周期 t 到指数 j 的一个任意映照，策略剖面｛周期 t 采取 $\alpha^{j(t)}$｝也是子博弈完美均衡。

从直观感觉，观察结果 8.1 是自然且合理的，如果想严格地证明结论 8.1 的正确性，在折扣因子 δ 满足 $0<\delta<1$ 条件时，仅需应用第五章中可观察行动多阶段博弈的无限形式的一步偏离准则。折扣因子的条件使无限重复博弈的盈利函数满足一步偏离准则中的在无穷处连续的要求。而在每一个周期中，在给定对手已取均衡

(博弈论(简本)) 七 1/2/1

策略的条件下，没有一个局中人会主动地偏离均衡策略。

其实，策略本身表明局中人 i 的对手的将来行动独立于局中人 i 现今的行动，因此对于局中人 i 来说，他的最佳反应是采取策略以使自己当今周期的盈利极大化，就是说，关于 a_{-i}^{t} 作出自己的最佳反应。这些策略实质上是第五章中讨论过的“开环”(open-loop)型策略。

根据上述可观察行动无限重复博弈的建模将证明有关一些事实，它们对于无限重复博弈的研究相当有用，因此不妨用定理的形式进行描述。

定理 8.2 在可观察行动无限重复博弈中，每一个真子博弈是策略同形的，或者说，在不同的真子博弈的策略空间之间存在着 1—1 对应。

证明：我们仅需证明在整体博弈与任何一个真子博弈的策略空间之间存在着 1－1 对应关系就足够了。

假定任意一个开始于历史 h^t 的真子博弈，不妨记作 $G(h^t, \infty)$。对于整体博弈的任何一个策略剖面 s（毫无疑问，这里不必讨论混合或行为策略），我们需要在 $G(h^t,\infty)$ 的策略剖面空间中找到与之对应的策略剖面 $\hat{s}$。最简单的办法就是按照 s 构造一个 $\hat{s}$，很容易地可以想到 $\hat{s}$ 对周期 $(t+\tau-1)$ 的处理就恰如 s 对于周期 τ 的处理，其中 $\tau=1,2,\cdots$。用数学形式表达，相当于令 $\hat{s}(h^t)=s(h^1)$（h^1 为整体博弈刚开始时的历史，常有 $h^1=\Phi$），$\hat{s}(h^t,a^t)=s(a^1)$，$\hat{s}(h^t, a^t,a^{t+1})=s(a^1,a^2)$，…，等等。反过来，对于给定历史 h^t 以及 $G(h^t, \infty)$ 上的策略剖面 $\hat{s}$，我们当然也可以由 $\hat{s}$ 出发构造整个博弈的策略剖面 s 与之对应。于是，得到了定理的证明。

注：如果 $G(\delta,\infty)$ 的策略剖面 s 与 $G(h^t,\infty)$ 的策略剖面 $\hat{s}$ 具有上述定理证明中的对应关系，那么，如果 s 是 $G(\delta,\infty)$ 的 Nash 均衡，则 $\hat{s}$ 必为真博弈 $G(h^t,\infty)$ 的 Nash 均衡，反之亦然。这样，我

们实际上已经证明了如下定理：

定理 8.3 对应于所有 Nash 均衡的持续盈利向量集合在重复博弈的每一个真子博弈(包括整体博弈)中均为相同。

2. 无限重复博弈的无名氏(Folk)定理

这是无限重复博弈理论的一个相当重要的定理，在专家们发表类似结论之前，早已在博弈论研究中广泛应用，因此称为无名氏定理。该定理指出，如果局中人充分耐心(例如 δ 充分接近 1)的话，任何可行的、个人理性的盈利可以通过均衡来实施。这在直观上似乎很容易被人们所接受，仿佛有点像：只要你有足够的耐心，不断地去奋斗争取，"面包总是会有的"。

现在我们正式地介绍无名氏定理，在此之前，必须弄清楚盈利的"可行性"与"个人理性"。先定义局中人 i 的保留效用(reservation utility)或最小最大值(minmax value)：

$$\underline{v}_i = \min_{\alpha_{-i}}\left[\max_{\alpha_i} g_i(\alpha_i, \alpha_{-i})\right] \tag{8.5}$$

(8.5)式里中括号的意思非常清楚，在给定对手的策略 α_{-i} 条件下，局中人 i 在阶段博弈中选择最佳反应 α_i 以使自己的静态盈利极大化。$\max\limits_{\alpha_i} g_i(\alpha_i, \alpha_{-i})$ 正是这个"条件极大化盈利值"。中括号外面的运算符号 $\min\limits_{\alpha_{-i}}$ 表明此时的主动权"掌握"在局中人 i 的对手手里。因为"条件极大化盈利值"显然依赖于"条件"，假如局中人 i 真的能正确地预测对手的策略 α_{-i} 并相应地作出最佳反应，随着对手取不同的 α_{-i}，将会有不同的条件极大盈利。对手通过选取特定的 α_{-i}，使得局中人 i 在这些不同的条件极大盈利中间只能得到最小的一个，(8.5)式正是表示这个值。令 m_{-i}^i 表示局中人 i 的对手使(8.5)式中这一最小值得以达到的策略。称 m_{-i}^i 为针对局中人 i 的最小最大剖面(minmax profile)。对应地，令 m_i^i 为此时局中人 i 的策略，因此有 $g_i(m_i^i, m_{-i}^i) = \underline{v}_i$。

例 8.1 设阶段博弈 G 的盈利矩阵如图 8.7 所示。

		局中人 2	
		L	R
局中人 1	U	$-2,2$	$1,-2$
	M	$1,-2$	$-2,2$
	D	$0,1$	$0,1$

图 8.7

先计算局中人 1 的 minmax 值，由定义，应该考虑所有的混合策略，故设局中人 2 以概率 q 取 L，以概率 $1-q$ 取 R，于是，局中人 1 关于纯策略 U、M、D 的盈利当然是 q 的函数：

$$u_1(U)=-2q+1-q=-3q+1$$

$$u_1(M)=q-2(1-q)=3q-2$$

$$u_1(D)=0$$

由于局中人 1 若取 D 的话，不管局中人 2 取 q 等于多少，其盈利总是等于 0。故局中人 1 的 minmax 值至少大于等于零。关键在于局中人 2 是否能取 q 使得局中人 1 的极大化盈利为 0。由于 $u_1(D)$中不出现 q，因此我们观察 $u_1(U)$与 $u_1(M)$，看看能否选择 q 使得这两个值均小于等于 0，显然当 $q=1/2$ 时，$u_1(U)=u_1(M)=-1/2<0$，这样，在局中人 2 取混合策略$(1/2,1/2)$时，局中人 1 的极大盈利通过取 D 达到 0，因此，$\underline{v}_1=0$，其实，由 $u_1(U)$与 $u_1(M)$的表达式可知，当 $q\in[1/3,2/3]$时，$\max(u_1(U),u_1(M))\leqslant 0$，因此局中人 2 取任何该范围的 q，都可以使局中人 1 的极大盈利通过取 D 达到$\underline{v}_1=0$。

再来寻求局中人 2 的 minmax 值$\underline{v}_2$。令局中人 1 以概率 p 取 U，以概率 r 取 M，混合策略为$(p,r,1-p-r)$，其中 $p+r\leqslant 1$。局中人 2 取 L 与 R 的盈利 $u_2(L)$，$u_2(R)$表示为 p，r 的函数：

$$u_2(L)=2(p-r)+(1-p-r) \tag{8.6}$$

$$u_2(R)=-2(p-r)+(1-p-r) \tag{8.7}$$

显然，

$$\underline{u_2}=\min_{p,r}\max[2(p-r)+(1-p-r),-2(p-r)+(1-p-r)] \tag{8.8}$$

也许读者会提出疑问，根据(8.5)式，应对局中人 2 的所有混合策略求给定条件(p,r)下盈利的极大值，而公式(8.8)式仅对两个纯策略盈利求极大值。是否会发生差错？事实上，对局中人 2 的任何混合策略$(q,1-q)$，相应的盈利是$u_2(L)$与$u_2(R)$的线性组合，只需将概率 1 加权于$u_2(L)$与$u_2(R)$之间最大者自然获取局中人 2 的极大盈利。这一思路有利于我们去求 minmax 值。现在先求(8.8)式中 max 部分的表达式：

$$\max(u_2(L),u_2(R))=\begin{cases}2(p-r)+(1-p-r)=1+p-3r & \text{当 } p>r\\ -2(p-r)+(1-p-r)=1-3p+r & \text{当 } p<r\\ 1-p-r & \text{当 } p=r\end{cases} \tag{8.9}$$

当 $p>r$ 时，由于 $0\leqslant p+r\leqslant 1$，因此必有 $r<1/2$，于是

$$1+p-3r=1+(p-r)-2r>0 \tag{8.10}$$

同理，当 $p<r$ 时，也有

$$1-3p+r=1-2p+(r-p)>0 \tag{8.11}$$

当 $p=r$ 时，总有 $0\leqslant p=r\leqslant 1/2$，因此

$$1-p-r\geqslant 0 \tag{8.12}$$

无论(p,r)如何取，总有 $\max(u_2(L),u_2(R))\geqslant 0$，除非 $p=r=1/2$，则有 $\max(u_2(L),u_2(R))=0$。据以上分析，我们断言，当局中人 1 取混合策略(1/2,1/2,0)时。局中人 2 无论取什么样的混合策略，总达到 minmax 值$\underline{v_2}=0$。注意到这样的事实，为使$\underline{v_2}=0$，局中人 1 的策略是唯一确定的，这与刚才所述，为使达到$\underline{v_1}=0$，局中人 2 的混合策略有一定的范围可取。两者之间存在明显的不同。

如果将公式(8.5)改为只考虑纯策略空间，以图 8.7 为例，读

者不难发现此时 minmax 值发生了改变。

为什么我们称 minmax 值为保留效用呢？这从下面的结果可以得到启发：

观察结果 8.2 在任何静态均衡和在重复博弈的所有 Nash 均衡中，局中人 i 的盈利至少为$\underline{v}_i$，不管折扣因子如何取值。

证明：在任何静态均衡中的结论是不言而喻的，因为在均衡 α 中，α_i 是对 α_{-i}的最佳反应，即在给定 α_{-i}下，$g_i(\alpha_i, \alpha_{-i})$达到极大值，由(8.5)式，它至少不会小于$\underline{v}_i$。

现在我们考虑重复博弈的任意一个 Nash 均衡 $\hat{\sigma}$。

局中人 i 的一个可行(或称行得通)的策略有可能是比较缺乏远见的策略之一，它选择每一周期的行动 $a_i(h^t)$以极大化 $g_i(a_i, \hat{\sigma}_{-i}(h^t))$的期望值。这个策略也许不是最佳的，因为它忽略了局中人 i 的将来行动依赖于他今天的行动的可能性。不管怎样，在每一个周期 t，必有

$$g_i(a_i(h^t), \hat{\sigma}_{-i}(h^t)) \geqslant \underline{v}_i \tag{8.13}$$

即可行策略尽管缺乏远见，但在每个周期为局中人 i 产生的盈利至少是$\underline{v}_i$。在无限重复博弈的 Nash 均衡 $\hat{\sigma}$ 中，给定对手的策略 $\hat{\sigma}_{-i}$，从整体来说，局中人 i 取 $\hat{\sigma}_i$ 的盈利当然不小于可行策略的盈利，而任何可行策略的盈利函数为

$$(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}g_i(a_i(h^t), \hat{\sigma}_{-i}(h^t)) \geqslant (1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}\underline{v}_i \geqslant \underline{v}_i \tag{8.14}$$

(8.14)式与 δ 无关(只要 $\delta \neq 1$)。它蕴含着$\underline{v}_i$ 是局中人 i 的均衡盈利的下界。

由观察结果 8.2，我们可以预先知道。无限重复博弈的任何均衡对于任何一个局中人来说，所带来的盈利不会少于他的 minmax 值。因此称局中人 i 的 minmax 值为他的保留效用完全有其实践上的意义。

基于 minmax 值的引入以及观察结果 8.2，我们引进关于“可行盈利”与“个体理性盈利”这两个概念。

既然已经知道在任何静态均衡和在重复博弈中的 Nash 均衡里局中人 i 的盈利至少为$\underline{v}_i$，又知道均衡是理性人理性博弈的合理结局，因此，作为盈利，如果称为个体理性的，则要求它大于$\underline{v}_i$。

接下来考虑可行盈利(或效用)。这是一个在无名氏定理中将涉及到的重要概念。

假设某人有 A、B 两个策略，各具有盈利(或期望盈利)1 与 2。他采取的任何混合策略$(p,1-p)$，其期望盈利应为

$$1\cdot p+2\cdot(1-p)$$

这个量当然位于 1 与 2 之间。人们自然会设想，对于任何实数 $R\in[1,2]$，只要该局中人采取适当的混合策略，相应的期望盈利一定可以达到 R。或者说，R 作为策略 A 与 B 相应盈利的凸组合，是个“行得通”的盈利。现在将这种思想移植到两人有限博弈中，所有可能结局的盈利向量的凸组合似乎也可以说是“行得通的”或者“可行的”盈利向量。然而事情并没有那么简单，因为一般的有限阶段博弈中可行的(就是说可达到的)盈利集合未必是凸的。以博弈论中一个著名的“性别战”(battle of the sexes)为例来说明这个问题。

夫妻俩希望一同观看娱乐活动，然而在“看什么”问题上产生不同意见：丈夫喜欢看足球而妻子喜欢看芭蕾舞剧。如果两人参加同一项活动，那么喜欢该项活动的一方获益 2 而另一方仅获益 1；如果双方未能达成一致意见而赌气待在家中或者干脆各人看自己喜欢的，则获益均为 0，博弈的盈利矩阵如图 8.8 所示。

该博弈有两个纯策略均衡(B,B)、(F,F)，各具有盈利向量$(1,2)$与$(2,1)$。非均衡的结局具有盈利$(0,0)$。如果对(B,B)与(F,F)的盈利向量各赋予 1/2 权，则得

$$1/2\times(1,2)+1/2\times(2,1)=(3/2,3/2)$$

		妻子	
		B(芭蕾)	F(足球)
丈夫	B	1,2	0,0
	F	0,0	2,1

图 8.8

那么丈夫与妻子是否可以通过独立地选择自己的混合策略从而达到(3/2,3/2)呢?

不妨令丈夫以概率 p 选择 B,妻子独立地以概率 q 选择 B,那么丈夫(局中人 1)与妻子(局中人 2)各自的期望获益如下并令它们均等于 3/2

$$g_1 = pq + 2(1-p)(1-q) = 3/2 \tag{8.15}$$

$$g_2 = 2pq + (1-p)(1-q) = 3/2 \tag{8.16}$$

比较(8.15)式与(8.16)式可得

$$pq = (1-p)(1-q) = 1-p-q+pq \Rightarrow p+q=1 \tag{8.17}$$

以(8.17)式代入(8.15)式得

$$p(1-q) + 2p(1-p) = 3/2 \Rightarrow p(1-p) = 1/2 \Rightarrow 2p^2 - 2p + 1 = 0 \tag{8.18}$$

(8.18)式显然无实数解。这表明性别博弈不可能通过局中人的独立选取混合策略而达到(3/2,3/2),或者说,该盈利向量在性别博弈中是不可行的(或行不通的)。因此在"性别战"这样的阶段博弈中,可行盈利集合不是凸集。

设想重复"性别博弈",夫妻俩人常常商量今晚是看芭蕾舞演出还是欣赏足球比赛。由于阶段博弈具有多重 Nash 均衡,根据观察结果 8.1,存在着多重子博弈完美均衡结局,不管这样的子博弈完美的具体形式如何,只要折扣因子 δ 充分地小,此时局中人表现得相当无耐心,相应的盈利函数基本上就是第一周期博弈的盈利。于是在折扣因子很小的重复博弈中,看来可行盈利向量集也未必

是凸集。但是，如果在重复博弈中，δ 相当接近于 1，或者说，局中人相当有耐心，此时通过随时间变化来确定途径的方法可以得到纯策略剖面向量的任何凸组合。试看一个简单的例子，仍以性别战作为阶段博弈，从第一周期开始凡奇数周期总取 (F,F)；凡偶数周期总取 (B,B)。由于图 8.8 中盈利周期矩阵关于两人局中人是对称的，我们仅需求丈夫的盈利或效用函数：

$$
\begin{aligned}
u_1 &= (1-\delta)\{2+\delta+2\delta^2+\delta^3+\cdots\} \\
&= (1-\delta)\{2(1+\delta^2+\cdots)+\delta(1+\delta^2+\cdots)\} \\
&= (1-\delta)\left\{\frac{2+\delta}{1-\delta^2}\right\} \\
&= \frac{2+\delta}{1+\delta} \qquad (8.19)
\end{aligned}
$$

令 $\delta\to 1$，则 $u_1\to 3/2$。同样，妻子的平均盈利 u_2 当 δ 充分接近于 1 时也趋于 3/2。

这个例子已经开始接近于无名氏定理的内容，但是，诸如时间变化路径这样的方法最好避免"必须使用"。Fudenberg 与 Maskin 于 1986 年通过假设在每个周期的开头所有局中人观察到共同随机化装置的结局，从而凸化阶段博弈的可行盈利向量。所谓共同随机化装置，就是我们在第二章中讨论相关均衡时所提及的，譬如掷一颗骰子，通过看它呈现的结果（点数，或单点数，或双点数等）确定局中人的行动。这种共同观察到的随机化装置使得局中人执行相关的策略行动，而正是策略行动的相关性导致纯策略盈利向量的凸组合可以达到。为使读者对这一点有直观上的了解，再一次研究性别博弈。我们设计的共同观察的随机化装置是——抛一枚均匀的钱币，如果出现正面，夫妻双双欣赏芭蕾舞剧；倘若出现反面，那么就可以在足球场上见到他俩的身影。读者不难立即计算得到，执行如此确定的相关策略行动，丈夫与妻子的期望盈利向量恰好是(3/2,3/2)。读者只要举一反三，如果阶段博弈有有限个结局伴随着有限个盈利向量，对于这些盈利向量的凸组合，我们可以设法

设计一个随机化装置，使得由此产生相关策略行动可以达到这个凸组合盈利向量。

看来，借助于共同随机化假设（或装置），我们可以将有限阶段博弈中纯策略剖面盈利向量的任何凸组合视作可行盈利向量了。正式地，令$\{w^1,\cdots,w^t,\cdots\}$为一系列来自$[0,1]$上均匀分布的独立抽样，并假定在周期$t$的开始时局中人观察到$w^t$。现在周期$t$的历史应当记作

$$h^t=(a^1,\cdots,a^{t-1},w^1,\cdots,w^t)$$

局中人i的纯策略s_i于是为一系列从历史h^t到A_i中的映射s_i^t。此时，对于任意的折扣因子δ，可行盈利集合为

$$V=\text{集合}\{v|\exists\ a\in A,\ \text{使}\ g(a)=v\}\text{的凸包} \tag{8.20}$$

v为盈利向量，(8.20)式中花括号内的v是指阶段博弈中纯策略向量所能达到的那些盈利向量v，基于这些v的凸包即为可行盈利集。

回到性别博弈，$v=(v_1,v_2)=(1,2)$、$(2,1)$和$(0,0)$。由这三个盈利向量构成的凸包V见图8.9阴影部分。

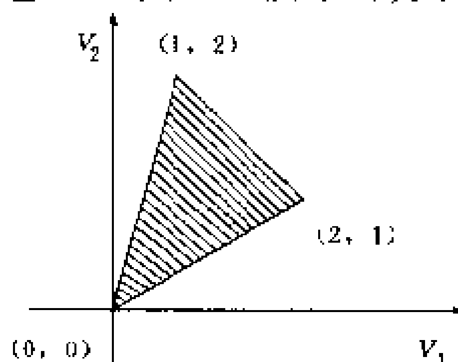


图8.9

如果对图8.7中的阶段博弈求可行盈利向量集，此时v包含$(-2,2)$、$(1,-2)$和$(0,1)$三个向量，由它们构成的凸包见图8.10。

凸包是指图中大三角形内部再加上边界，而图中阴影部分中的盈利显然Pareto优于minmax盈利$(0,0)$。因此根据定义，该集

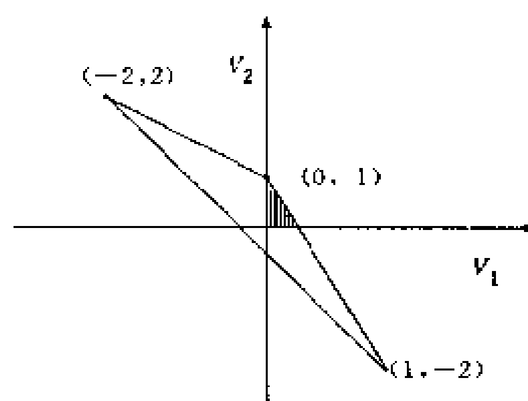


图8.10

合$\{v\in V|v_i>\underline{v}_i,i=1,\cdots,n\}$实际上是阶段博弈的可行且严格个体理性盈利向量集合。

一旦可行,个体理性的盈利向量集被定义,我们就可以叙述并论证无限重复博弈的无名氏定理。

定理8.4 (无名氏定理)对于每一个可行的盈利(效用)向量v,其中对所有局中人i成立$v_i>\underline{v}_i$,存在一个$\underline{\delta}(\underline{\delta}<1)$,使得对一切$\delta\in(\underline{\delta},1)$,存在$G(\delta,\infty)$的一个Nash均衡具有盈利$v$。

证明:首先假设一种特殊的情况:对于v,存在阶段博弈的一个纯策略行动剖面a达到该可行盈利,即$g(a)=v$。现在考虑局中人$i(i=1,2,\cdots,n)$的如下策略:

"在第一周期取a_i,并且继续采取行动a_i,只要

(1)在以前的周期中已实施的行动是a;或者

(2)在以前的周期中已实施的行动有两个以上的分量不同于a。

如果在以前某个周期中局中人i是唯一不遵循行动剖面a的人,那么每个局中人j在博弈的其余部分取m_j^i。"

局中人i能否通过偏离这个策略剖面而因此获益呢?根据以上策略剖面的定义,局中人i基本上一直取行动a_i,只有两种情况他可能偏离而不取a_i:

(1)只有一个其他局中人 j 偏离 a 剖面时，局中人 i 将与众人一起采取惩罚局中人 j 的行动；

(2)局中人 i 在某周期独自偏离而不取 a_i，使他自己在该周期获得更大好处。

问题是针对(2)而言，因为情况(1)是惩罚他人而用。属"不得已而为之"。

在局中人 i 独立主动偏离的周期里，他至多获益 $\max_a g_i(a)$（注：这里的 a 是泛指所有的纯策略行动剖面，与前面的 $g(a)=v$ 的 a 有所不同），可是，这种偏离将招致惩罚，在他第一次偏离后的所有周期中他将至多获得他的 minmax 值$\underline{v}_i$。于是，如果局中人 i 在周期 t 第一次偏离，他至多获益为

$$(1-\delta)(v_i+\delta v_i+\cdots+\delta^{t-1}v_i+\delta^t \max_a g_i(a)+\delta^{t-1}\underline{v}_i+\delta^{t+2}\underline{v}_i+\cdots)$$

$$=(1-\delta^t)v_i+(1-\delta)\delta^t \max_a g_i(a)+\delta^{t+1}\underline{v}_i \tag{8.21}$$

令(8.21)式小于 v_i(即考虑局中人 i 不愿偏离的情况)，并使得其中 $\delta=\delta_i$，可得

$$v_i>(1-\delta_i)\max_a g_i(a)+\delta_i\,\underline{v}_i \tag{8.22}$$

由(8.22)式，我们仅需取

$$\delta_i=(\max_a g_i(a)-v_i)/(\max_a g_i(a)-\underline{v}_i) \tag{8.23}$$

(8.23)式中，由于 v 是可行且个体理化盈利，必有 $v_i>\underline{v}_i$，因此

$$\max_a g_i(a)\geqslant v_i>\underline{v}_i \tag{8.24}$$

故分母不可能等于 0，而若 $v_i=\max_a g_i(a)$，则局中人 i 不存在偏离行为，故 $0<\underline{\delta}_i<1$。令

$$\underline{\delta}=\max_i(\underline{\delta}_i) \tag{8.25}$$

实际上就 $g(a)=v$(a 为纯策略行动剖面)类型的 v 已经证明了无名氏定理。因为我们找到了$\underline{\delta}$ 及一个策略剖面，使得任何一个局中人都不会主动偏离，这个策略剖面当然是 Nash 均衡。注意我们在

证明中，决定局中人 i 是否在周期 t 偏离时，实际上对于局中人 i 的其他对手在该周期的偏离指派了 0 概率，这种做法符合 Nash 均衡的定义——只考虑单方面的偏离。

问题将转到"可行个体理性盈利向量 v 不可能由完全纯策略行动剖面产生"上来。设想有一个行动剖面 a，以及相应的共同随机化装置 $a(w)$，由它们产生期望值为 v 的盈利。注意到上面这句话的重要性，$a(w)$ 产生的盈利的期望值为 v，对局中人 i 来说，期望盈利就是 v_i。因此可以说，局中人 i 在每个周期得到的盈利不可能恰好都是 v_i，总是在有些周期高一些，有些周期低一些。局中人 i 不会在高出 v_i 的那些周期偏离 $a(w)$（这里策略剖面的叙述仍同前面 a 为纯策略时一样，只要将 a 改为 $a(w)$），只有在 $g_i(a(w))$ 偏低的局期，他才有可能受到高盈利的诱惑而偏离 $a(w)$。假设在周期 t，局中人 i 第一次偏离，我们仅需计算从 t 开始的持续盈利即可。像前面一样，在 t 以后的周期他将受到惩罚，每个局期的获益至多为$\underline{v}_i$，故持续盈利不超过

$$
\begin{aligned}
&(1-\delta)(\max_a g_i(a)+\delta\,\underline{v}_i+\delta^2\,\underline{v}_i+\cdots)\\
&=(1-\delta)\max_a g_i(a)+\delta\underline{v}_i \qquad (8.26)
\end{aligned}
$$

尽管在周期 t 时 $g_i(a(w))$ 较小，如果局中人 i 不偏离 $a(w)$，那么他的持续盈利应该是多少呢？由于从周期$(t+1)$开始，他在每个周期的平均获益为 v_i，故不发生偏离的持续盈利为

$$
\begin{aligned}
&(1-\delta)(g_i(a(w))+\delta v_i+\delta^2 v_i+\cdots)\\
&=(1-\delta)g_i(a(w))+\delta v_i\\
&\geqslant(1-\delta)\min_a g_i(a)+\delta v_i \qquad (8.27)
\end{aligned}
$$

之所以在(8.27)式中应用一个显然的不等式，原因在于避开随机化装置的影响，定理中的$\underline{\delta}$ 不能随该装置而随机变化。欲使局中人 i 不偏离策略剖面，只要 δ_i 满足

$$
(1-\delta_i)\min_a g_i(a)+\delta_i v_i>(1-\delta_i)\max_a g_i(a)+\delta_i\underline{v}_i \qquad (8.28)
$$

就足够了，令

$$\underline{\delta}_i=\frac{(\max\limits_a g_i(a)-\min\limits_a g_i(a))}{[(v_i-\underline{v}_i)+\max\limits_a g_i(a)-\min\limits_a g_i(a)]} \tag{8.29}$$

只要 $\delta_i>\underline{\delta}_i$，(8.28)式就可满足。令

$$\underline{\delta}=\max(\underline{\delta}_1,\underline{\delta}_2,\cdots,\underline{\delta}_n) \tag{8.30}$$

构造的策略剖面在 $\delta\in(\underline{\delta},1)$时是一个 Nash 均衡，且具有可行的和个人理性的盈利向量 v(可能要借助于共同随机化装置)。比较(8.23)式与(8.29)式可知，(8.29)式中的$\underline{\delta}_i$ 一般大于(8.23)式中的$\underline{\delta}_i$，相应地$\underline{\delta}$当然也有此特点。因此，在通过共同随机化装置而达到期望盈利 v 的情况下，通常折扣因子下界$\underline{\delta}$ 要求稍大一些才足以保证策略剖面是 Nash 均衡。

在无名氏定理中用于证明的策略下，单方面偏离将会招致冷酷无情的惩罚。对于惩罚者来说，这样的惩罚可能是昂贵的。譬如在产量设置方面的 Cournot 双寡竞争，如果重复地实施(实际情况也的确如此)，惩罚局中人 i 的策略——由于惩罚而使局中人 i 仅获益 minmax 值，故简单地称为 minmax 策略——要求局中人 i 对手生产许多产品，使得价格落到局中人 i 的平均成本以下，但是这样做的结果也可能使价格落到惩罚者自己的平均成本之下，由于 minmax 惩罚可能是昂贵的，问题呈现为，局中人 i 是否会因为担心他的对手做出冷酷无情的惩罚这一反应而吓得不去作一次性的有利于自己的偏离。正式一些讲，关键在于我们不知道用来证明无名氏定理的策略剖面是否为子博弈完美。这就提出了如下问题：无名氏定理的结论能否应用于完美均衡的盈利向量。这个问题在 1971 年由 Friedman 作了回答。

定理 8.5 “Nash 威胁”无名氏定理(Friedman 1971)：设 a^* 为阶段博弈的 Nash 均衡，相应的盈利向量记作 e。那么对任意 $v\in V$，其中 $v_i>e_i$ 对所有局中人 i 均成立，存在一个$\underline{\delta}$，使得对所有

$\delta>\underline{\delta}$，存在无限重复博弈 $G(\delta,\infty)$ 的子博弈完美均衡，其盈利为 v。

在尚未证明定理 8.5 之前，先比较一下无名氏定理与"Nash 威胁"无名氏定理，前者讨论的 v 属于可行且个人理性盈利向量集合，后者讨论的 v 属于可行且 Pareto 优于均衡盈利的盈利向量集合。由于 Nash 均衡盈利对于各局中人来说，不会少于 minmax 值，因此前者的 v 包含了后者 v 的范围。于是满足无名氏定理结论的策略剖面自然地包含了满足"Nash 威胁"无名氏定理结论的策略剖面。这样的比较使人们对定理 8.5 的结论感到非常自然。由此启发，既然无名氏定理的证明中以 minmax 值来威胁局中人以致他不敢偏离。那么可以设想在定理 8.5 中也可以用 Nash 均衡值来威胁局中人，这就是定理 8.5 有时候称作"Nash 威胁"无名氏定理的由来。

现在我们试图证明定理 8.5。

证明：假设存在 $\hat{a}$ 使得 $g(\hat{a})=v$，考虑如下策略剖面：

在第一周期，每个局中人 i 取 $\hat{a}_i$。在以后的周期里，只要在所有过去的局期中实施的行动是 $\hat{a}$，那么局中人 i 就继续取 $\hat{a}_i$。如果至少有一个局中人不按 $\hat{a}$ 行动，那么在博弈的其余局期每个局中人 i 采取行动 α_i^*。

先证明这个策略剖面（对充分大的 δ）是 Nash 均衡。

如果局中人 i 在周期 t 偏离 $\hat{a}_i$，那么从 t 开始的持续盈利应当不会超过

$$(1-\delta)\max_a g_i(a)+\delta e_i$$

上式当 $\delta=1$ 时，严格地小于 v_i（因为 $v_i>e_i$），因此当 δ 小于 1 但很接近于 1 时，成立不等式

$$(1-\delta)\max_a g_i(a)+\delta e_i<v_i \qquad (8.31)$$

这表明，当 δ 充分大地接近于 1 时，局中人 i 不取 $\hat{a}_i$ 则会在盈利方面造成损失，因此策略剖面是 Nash 均衡。

现在需要证明这个策略剖面是子博弈完美的。由定理 8.2，$G(\delta,\infty)$的每一个真子博弈等同于 $G(\delta,\infty)$本身。在上述 Nash 均衡中，子博弈可以群分为两类：①所有早先周期的行动为 $\hat{a}$；②至少有一个早先周期局中人的行动不同于 $\hat{a}$。对于类型①的子博弈，同样的论证方法告诉我们所构造的策略剖面是 Nash 均衡。对于类型②的子博弈，显然局中人的策略行动简单地为一直重复着阶段博弈 Nash 均衡 α^*。因此从总体上它是该子博弈的 Nash 均衡。既然策略剖面在所有子博弈上均为 Nash 均衡。它自然是子博弈完美均衡。

又到了需要考虑不存在 $\hat{a}$ 使得 $g(\hat{a})=v$ 的情况，因此又得借助于共同观察的随机化装置，以下的叙述完全雷同于定理 8.4 的相应部分，只不过在检验它是子博弈完美时又要重复刚才的一段叙述，读者不妨可以自己试作练习。

在证明无名氏定理和 Friedman 定理时，我们构造的策略有一个共同的特点，任何一个局中人的一次性不合作（偏离 a 或 $\hat{a}$）将触发局中人永远地不合作的开关，我们称这样的策略为触发策略(trigger strategies)。

我们在比较无名氏定理与 Friedman 定理时曾提及一种直觉：可行且个人理性的向量集合包含了可行且优于 Nash 均衡的盈利向量集合，因此达到前者 v 的策略剖面一定包含了达到后者 v 的策略剖面。无名氏定理结论中的结局为 Nash 均衡，而 Friedman 定理中的结局是子博弈完美的。这两种结果似乎体现出了这一直觉。人们不禁会问，在无限重复博弈中，子博弈完美均衡能达到的是否一定局限于可行且优于 Nash 均衡的盈利向量，换句话问，可行且个人理性的盈利向量能否在一个子博弈完美均衡中达到。这是 Friedman 定理没有回答的问题。事实上，如果我们认真且仔细地研究这两个定理，立即发现定理的结论与$\underline{\delta}$ 有着强烈的关系，譬

如，我们已经指出，Friedman 定理中的折扣因子比起无名氏定理中的折扣因子要稍微大一些。其实，定理的证明过程隐隐地告诉我们，δ 越大越接近于 1，则情况似乎越好。Aumann 与 Shapley (1976)，Rubinstein(1979)，Fudenberg 与 Maskin(1986)证明了如下事实：对任意可行的且个人理性盈利向量，存在折扣因子的一个范围，使得盈利向量可以在一个子博弈完美均衡中得到。我们先介绍 1976 年 Aumann 与 Shapley 的研究成果，在他们的定理中阶段博弈效用系列的计算采用了时间平均准则(见(8.4)式)，这是局中人完全具有耐心时数学建模的盈利函数计算。此时 $\delta=1$，局中人对盈利的时间性不关心，而且对任意有限个周期的盈利同样表示出不关心。

定理 8.6 (Aumann & Shapley 1976)如果局中人利用时间平均准则来估算阶段博弈效用系列，那么对任意可行盈利向量 $v\in V$，其中 $v_i>\underline{v}_i$ 对所有局中人 i 成立，存在一个子博弈完美均衡，其盈利向量为 v。

证明：证明的手段还是构造一个策略剖面：

“以‘合作状态’开始，在这个状态中，借助于共同随机装置，采用达到 v 的策略 p，只要不存在偏离，则每个局中人均留在该状态中。如果某局中人 i 偏离合作状态，那么所有局中人取 minmax 策略 $m^i=(m_i^i,m_{-i}^i)N$ 个周期，其中 N 的选择满足如下不等式：

$$\max_a g_i(a)+N\underline{v}_i<\min_a g_i(a)+Nv_i \qquad (8.32)$$

经过这 N 个周期之后，重新回到合作状态，不管在这 N 个周期中是否存在任何关于 m^i 的偏离。”

根据策略剖面的定义，它具有盈利向量 v。为证实它是子博弈完美均衡，仅需核实在任何子博弈中不存在改善局中人盈利的策略。任意一个局中人 i 如果在某周期偏离合作状态，那么此时他至多获得 $\max_a g_i(a)$，但是在随后的 N 个周期将受到惩罚，共仅获益

$N\underline{v}_i$。因此，在这 $N+1$ 个周期内共获益 $\max_a g_i(a)+N\underline{v}_i$。局中人 i 为什么在该周期可能偏离呢？理由在于共同随机装置产生的策略使他在该周期的盈利可能低于 v_i（因为策略剖面仅保证每周期的期望盈利为 v_i）。但其盈利至少不会少于 $\min_a g_i(a)$，而在随后的 N 个周期期望盈利为 Nv_i。条件(8.32)式所定义得到的 N 实际上保证了局中人从偏离合作状态的任意获益在惩罚状态中被冲掉了。因此，没有一个系列的有限或无限偏离能让局中人 i 的平均盈利得到增加而超过 v_i。况且，对于惩罚者（任何局中人 j）来说，即使对偏离者局中人 i 实施 minmax 惩罚，从每周期的盈利而言代价是较大的，但是由于盈利的计算是按时间平均准则进行的，任何有限次（即 N 次）这样的损失对总平均盈利无甚损失。于是，在局中人 i 受到惩罚的子博弈中，局中人 j 的平均盈利仍为 v_j，这样没有一个局中人 j 可以通过在任何子博弈的偏离（即不去惩罚局中人 i）而获益。综上所述，所构造的策略剖面是子博弈完美均衡。

Aumann 与 Shapley 定理的奥妙之处在于使用时间平均准则计算得失，因为从总体来讲，这种计算方法使每一个局中人不在乎有限个周期内因为惩罚他人而蒙受的损失。前面已经指出，在该种情况下，$\delta=1$——局中人完全地有耐心。倘若 δ 非常接近于 1，现在局中人就会关心这样的事：在有限个周期内用 minmax 策略惩罚偏离者而招致自己的损失。如果这种惩罚使自己在有限个周期内有可能盈利低于自身的 minmax 值，这一来就有点得不偿失，因为$\delta\neq 1$时这段时间的盈利对总盈利有所影响。从博弈论观点，敢于这样地去惩罚他人的局中人至少不是理性的。于是人们感兴趣于如下的研究：当存在折扣因子时，是否存在子博弈完美均衡使得其盈利属于可行且个人理性盈利集。Fudenberg 与 Maskin 于 1986 年从另一种角度探讨了这个问题并得到了类似无名氏定理的结果，主要手段是考虑不同类型的策略——以引导局中人 i 的对手

们用 minmax 策略对付局中人 i。如果他们不用 minmax 策略对付局中人 i 的话，将不是用"惩罚"来加以威胁，而是宣布给予"奖励"以刺激他们采用 minmax 策略去惩罚局中人 i。在设计一个策略剖面以对惩罚偏离者提供奖励时，人们必须当心不要同时去奖励原来的违规者，否则奖罚不明将引起对违规偏离行为的鼓励。Fudenberg 与 Maskin 在解决这个问题的过程中引进了"满维"条件。

定理 8.7 (Fudenberg & Maskin)假定可行盈利集 V 的维数等于博弈局中人的个数。那么，对于可行且个人理性盈利向量集合中的任何盈利向量 v，存在一个折扣因子 $\underline{\delta}<1$，使得对一切 $\delta\in(\underline{\delta}, 1)$，存在一个 $G(\delta,\infty)$ 的子博弈完美均衡具有该盈利向量 v。

本定理的证明较长，数学味道也稍微浓了一点，我们不准备在这里作详细介绍，对此感兴趣的、想了解所以然的读者不妨参看 Fudenberg 与 Tirole 的"Game Thery"一书中第 157～159 页。我们仅对本定理特有的"满秩"条件稍作解释。最好的解释也许是举出一个不满维的可行盈利集合的例子来。回忆猜谜游戏博弈，共有两个局中人，各有两个纯策略：伸出 1 根手指或 2 根手指。四个结局却只有两个可分辨的盈利向量：(1，－1)与(－1，1)。显然连接(1，－1)与(－1，1)两点的线段是可行盈利集合，它是一维的，其维数不等于局中人个数—— 2。然而，读者不难验证，以猜谜游戏作为阶段博弈的重复博弈，子博弈完美的无名氏定理是成立的。可见，定理 8.7 中可行盈利集合的"满维"要求是充分的但不是必要的，1990 年 Abreu 与 Dutta 就将满维条件减弱为"V 的维数＝局中人的个数－1"。

3. 均衡集的特性

我们所提到的无名氏定理的各种形式均强调了"折扣因子 δ 充分地接近于 1"这个条件。另外一个问题也是饶有兴趣的：δ 不充分地接近于 1。情况又会怎样呢？或者说，当 δ 为固定值时(当然它

不是太小)，无名氏定理的结论又会是怎样？从前面叙述的几个定理可知，如果 δ 是个比较大的但却是固定的折扣因子时，有可能存在许多均衡。例如，在一定的条件下，对于同一个 v 可以取到一个 δ，存在两个子博弈完美均衡达到 v。当局中人 i 发生偏离，一个是对手使用 Nash 威胁，另一个则是使用 minmax 威胁，这两种威胁都旨在阻止局中人 i 的违规及偏离。显然，同样是子博弈完美，后一种威胁比前一种有效。人们乐于考虑这样的策略结构：它使得对局中人 i 的任何偏离的惩罚，是通过转向一个使他的盈利或效用达到最低的完美均衡来完成，因为这样更有效一些。问题在于这种最坏的均衡是否存在，下面两个定理回答了这个问题。

定理 8.8 (Fudenberg & Levine 1983)如果阶段博弈具有有限个纯策略行动，那么对每一个局中人 i 存在一个最坏的子博弈完美均衡 $\underline{w}(i)$。

证明：注意到阶段博弈只有有限个纯策略行动，因此每个周期的盈利自然为一致有界，再注意到无限博弈的盈利计算中存在着折扣因子 $\delta<1$，因此博弈及盈利满足在无穷远处连续的条件(见第五章，在叙述与证明无限水平博弈的一步偏离准则时曾引入"在无穷远处连续"概念)。因此，子博弈完美均衡集在策略乘积拓扑空间中为紧致集，这些策略的盈利在该拓扑空间也为连续。因此，对于每一个局中人来说存在着最坏与最好的子博弈完美均衡。

注 1：对于经济管理类读者来说也许对拓扑空间并不熟悉，不妨"粗糙地"想象，博弈在无穷远处连续，子博弈完美均衡集构成紧集且该紧集上的盈利函数也为连续，再根据紧集上连续函数有极大(与极小)值这个众所周知的事实，不难明白存在着最好与最坏的子博弈完美均衡。当然，这样想象只能作为"帮助理解"而不能代替证明。

注 2：由于无限重复博弈均衡集的平稳性，在任意子博弈中，

$\underline{w}(i)$也是最坏的均衡。

定理 8.9 (Abreu 1988)在阶段博弈中，如果

(1)每一个局中人的行动空间是有限维欧氏空间ロ的紧子集；

(2)每一个局中人的盈利是连续的；

(3)存在一个静态的纯策略均衡。

那么，对于每一个局中人，存在一个最坏的子博弈完美均衡$\underline{w}(i)$。

证明：对于所有的纯策略子博弈完美均衡中局中人 i 的盈利，总存在一个下确界，从纯策略子博弈完美均衡集中总可以选出一个系列 $s^{i,k}(k=1,2,\cdots)$，当 $k\to\infty$时，$g_i(s^{i,k})\to y(i)$。对应于每一个纯策略子博弈完美均衡 $s^{i,k}$，记 $a^{i,k}$为相应的均衡路径：

$$a^{i,k}=\{a^{i,k}(1),\ a^{i,k}(2),\cdots,\ a^{i,k}(t),\cdots\}$$

已知阶段博弈的策略行动空间 A 是有限维欧氏空间中的紧子集，因此纯策略行动序列集$\{a^{i,k},k=1,2,\cdots\}$也是紧集，不妨令 $a^{i,\infty}$为该紧集的聚点。由前述事实，对应于 $a^{i,\infty}$的局中人 i 的盈利当然为 $y(i)$。

有了上述准备工作后，对固定的局中人 i，考虑如下构造的策略剖面：

以状态 I_i 开始。

状态 I_i：采取行动序列

$$a^{i,\infty}=\{a^{i,\infty}(1),\ a^{i,\infty}(2),\cdots\}$$

只要不存在关于该序列的单方面偏离的话。如果局中人 j 在 t 周期单方面地偏离，那么在$(t+1)$周期开始状态 I_j。即，在$(t+1)$周期取 $a^{j,\infty}(1)$，在$(t+2)$周期取 $a^{j,\infty}(2)$，…，一直下去。

上述策略剖面可以看作是让局中人 i 取得最可能的低盈利，如果有另外一个局中人 j 不这样去做，那么就惩罚局中人 j，让他从此倒霉。

显然，如果所有局中人都遵循这些策略的话，局中人 i 的盈利只能等于 $y(i)$。但这样的策略剖面是不是子博弈完美的呢？这就是我们关心的问题。

设想这些策略如果不是子博弈完美的，那么必定存在局中人 i 与局中人 j，行动 $\hat{a}_j$，以及某任意 $\varepsilon>0$，还有 τ，使得

$$(1-\delta)g_j(\hat{a}_j,a_{-j}^{i,\infty}(\tau))+\delta y(j)>(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}g_j(a^{i,\infty}(\tau+t))+3\varepsilon \tag{8.33}$$

(8.33)式的左端是在 τ 周期中，局中人 j 发生偏离 I_i 状态后的持续盈利，(8.33)式的右端去掉 3ε 部分表示在大家遵循 I_i 条件下局中人 j 不单方面发生偏离的持续盈利。如果策略不是子博弈完美，那么一定存在一个周期 τ，使局中人 j 偏离之后反而获益，即，偏离后的持续盈利会严格地大于不发生偏离的持续盈利，因此即使不发生单方而偏离的持续盈利加上 3ε（由于 ε 为任意小的正数，故 3ε 也为任意小的正数）后，不等的关系仍保持严格地成立。这正是(8.33)式所叙述的意思。

由于定理假设(2)，每一个局中人的盈利是连续的，并且已知 $a^{i,k}\to a^{i,\infty}$，因此对充分大的 k，成立

$$(1-\delta)g_j(\hat{a}_j,a_{-j}^{i,k}(\tau))+\delta y(j)\geqslant(1-\delta)g_j(\hat{a}_j,a_{-j}^{i,\infty}(\tau))+\delta y(j)-\varepsilon \tag{8.34}$$

$$(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}g_j(a^{i,k}(\tau+t))\leqslant(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}g_j(a^{i,\infty}(\tau+t))+\varepsilon \tag{8.35}$$

综合(8.33)式、(8.34)式与(8.35)式，可得

$$(1-\delta)g_j(\hat{a}_j,a_{-j}^{i,k}(\tau))+\delta y(j)>(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}g_j(a^{i,k}(\tau+t))+\varepsilon \tag{8.36}$$

注意到 $a^{i,k}$ 是对应于子博弈完美均衡 $s^{i,k}$ 的路径，设想局中人在 τ

周期之前均遵循 $s^{i,k}$，而到了 τ 周期，局中人 j 不取 $a_j^{i,k}(\tau)$ 而取 $\hat{a}_j$（偏离了 $s^{i,k}$），在 τ 周期之后局中人遵循某均衡策略，从 $(\tau+1)$ 周期开始的局中人 j 的持续盈利 $z_j(\tau,\hat{a}_j)$ 显然不会小于 $y(j)$，因此

$$(1-\delta)g_j(\hat{a}_j,a_{-j}^{i,k}(\tau))+\delta z_j(\tau,\hat{a}_j)\leqslant(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}g_j(a^{i,k}(\tau+t)) \tag{8.37}$$

(8.37)式的成立纯粹由于 $s^{i,k}$ 为子博弈完美的缘故。由于 $z_j(\tau,\hat{a}_j)\geqslant y(j)$，故有

$$(1-\delta)g_j(\hat{a}_j,a_{-j}^{i,k}(\tau))+\delta y(j)\leqslant(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}g_j(a^{i,k}(\tau+t)) \tag{8.38}$$

(8.38)式与(8.36)式的相互矛盾证明了构造的策略是子博弈完美均衡。

上面两条定理告诉我们，局中人 i 的最坏子博弈完美是存在的。还有一些类似的结果，本书不准备一一介绍。我们最感兴趣的是，如果周中人 i 存在最坏于博弈完美均衡，那么在构造的均衡中，一旦局中人 i 单方面偏离，其他局中人不是采取 Nash 威胁或者 minmax 威胁等惩罚措施，而是给予最严厉的威胁，以阻止这种“违规破坏”行为。这是在理论上行的通而且人人都能接受的想法，然而，寻求每一个局中人的最坏可能是相当复杂的。相对而言，在对称博弈中寻求最坏的强对称纯策略均衡要容易一些，特别地，如果存在产生任意低盈利的强对称策略时更是如此。所谓“强对称”，是指对所有的历史 h^t 及任何两个局中人 i 与 j，成立 $s_i(h^t)=s_j(h^t)$，由此，即使观察到的历史本身不对称，这两个局中人采取相同的策略。这里须强调的是，不同局中人采取相同策略，通常熟知的针锋相对(tit for tat)策略就可能不是强对称的，因为这种策略是局中人采用上一个周期对手所采取的行动。例如，在重复囚徒窘境中，第一周期的行动为(坦白，抗拒)，这作为第二周期博弈前的

历史是不对称的，倘若采取针锋相对策略，此时局中人 1 取“抗拒”，局中人 2 取“坦白”，行动并不一致。

策略剖面既然有“强对称”，想当然也有“弱对称”（有时简称“对称”），如果 $h_1^t=\widetilde{h}_2^t, h_2^t=\widetilde{h}_1^t$，那么成立

$$s_1(h_1^t, h_2^t)=s_2(\widetilde{h}_1^t, \widetilde{h}_2^t)$$

则称策略剖面在弱的意义下对称，事实上，读者不难从上面定义看出，所谓弱对称，就是如果将过去的历史置换一下，那么两个局中人的当前行动也就发生置换。

早在 1986 年，Abreu 就证明了在如下的对称博弈中，最坏强对称均衡非常容易刻画。该对称博弈的行动空间为实数区间，盈利函数连续且有上界，而且满足：

条件(a)：对称纯策略剖面 $\vec{a}$（即，每个局中人 i 均取 a 的策略剖面）的盈利函数是拟凹的（quasiconcave，读者倘若不熟悉这个概念，可见本书第三章），并且当 $a\to\infty$时无界地减少。

条件(b)：令$\vec{a}_{-i}$表示这样的策略剖面，其中局中人 i 的所有对手选择行动 a、偏离纯策略对称剖面 $\vec{a}$的极大盈利，$\max\limits_{a'_i} g_i(a'_i, \vec{a}_{-i})$，随着 a 的增加而减弱。

在对称设置产量博弈中，条件(b)是自然的，在那里，公司通过生产相当大量的产品（即 $a\to\infty$）使价格趋于 0。于是使得偏离的最佳盈利非常地小。

在强对称均衡定义中，我们强调在非均衡路径与均衡路径上同样地要求对称性，这样排除了许多可以用来实施对称均衡结局的非对称的惩罚。

定理 8.10 （Abreu 1986）对称博弈若满足条件(a)与(b)，令 e^* 和 e_* 表示在纯策略强对称均衡中每个局中人的最高与最低盈利，

(1)盈利 e_* 可以在具有如下形式的强对称策略的均衡中达

到："开始于状态 A，其中局中人采取行动 a_*，a_* 满足

$$(1-\delta)g(\vec{a}_*)+\delta e^*=e_* \tag{8.39}$$

如果有任何偏离，则继续状态 A。否则，转向具有盈利 e^* 的完美均衡（即状态 B）。"

(2)盈利 e^* 可以以如下策略达到："只要没有偏离，则取常数行动 a^*，如果存在任何偏离，则转向最坏强对称均衡。"其他可行盈利可以以类似的方式达到。

§8.3 无限重复博弈的若干例子

1. 无限重复囚徒窘境

在第五章中已经讨论过无限重复囚徒窘境。仍以图 5.1 为阶段博弈。观察结果 8.1 实际上指出第五章中该问题的第一个子博弈完美均衡。由于（坦白，坦白）（即（背叛，背叛））是阶段博弈唯一 Nash 静态均衡，因此每个囚徒在每一个周期"坦白"是子博弈完美均衡。

（抗拒，抗拒）（即（合作，合作））的盈利向量为(1,1)，显然大于(0,0)——静态均衡的盈利。由 Fredman 的完美无名氏定理，对于适当的$\underline{\delta}$，当 $\delta\in(\underline{\delta},1)$时，"每个局中人从一开始就一直采取策略'抗拒'，只要没有一个局中人偏离。倘有人偏离，则在以后永远取策略'坦白'"是子博弈完美，这里的$\underline{\delta}=\frac{1}{2}$，它的计算只需要根据公式

$$(1-\delta)\max_a g_i(a)+\underline{\delta}e_i=v_i$$

即可得到。在图 5.1 的博弈中，$e_i=0$，$v_i=1$，$\max_a g_i(a)=2$，故立得 $\underline{\delta}=\frac{1}{2}$。已知（合作，合作）是博弈的有效结局，在一次囚徒窘境中，有效结局不是均衡，但在无限重复囚徒窘境中从一开始取有效结

局可能构成子博弈完美均衡。这个特性是经济学家尤其感兴趣的。

2. Cournot 双寡垄断之间的合作(collusion)

回忆静态 Cournot 博弈，若市场总产量 $Q=q_1+q_2$，市场出清价 $P(Q)=a-Q(Q<a)$。设每个公司具有边际成本 c，但无固定成本。公司同时选择产量 q_1 与 q_2。在唯一的 Nash 均衡中，每个公司生产 Cournot 产量 $q_c=\frac{1}{3}(a-c)$。此时两个公司各得盈利 $\frac{1}{9}(a-c)^2$。但如果两个公司各生产垄断产品的一半：$\frac{q_m}{2}=\frac{1}{4}(a-c)$，各获得盈利 $\frac{1}{8}(a-c)^2$，显然此时这两个公司的日子要好过得多。

考虑以 Cournot 竞争作为无限重复的阶段博弈，折扣因子为 δ，构造如下触发策略：

在第一个周期各生产一半垄断产品 $\frac{q_m}{2}$，在每一个周期 t，只要不发生偏离，则继续各生产 $\frac{q_m}{2}$；如果发生单方面偏离，则转向生产 Cournot 产量 q_c。

应用 Nash 威胁(完美)无名氏定理，对于适当大的 δ，该触发策略是无限重复博弈的子博弈完美 Nash 均衡。这个 $\delta\in(\underline{\delta},1)$ 中的 $\underline{\delta}$ 还是通过公式(不妨令 $i=1$)

$$(1-\delta)\max_a g_1(a)+\delta e_1=v_1$$

求得。这里 $v_1=\frac{1}{8}(a-c)^2$，$e_1=\frac{1}{9}(a-c)^2$，关键求局中人 1 若在某周期 t 偏离时可能的最大盈利。在该周期 t，局中人 2 仍取 $q_2=\frac{q_m}{2}=\frac{1}{4}(a-c)$，因此需要解

$$\max_{q_1}\left[a-c-\frac{1}{4}(a-c)-q_1\right]q_1 \tag{8.40}$$

易得 $q_1=\frac{3}{8}(a-c)$，代入局中人 1 的盈利函数(q_2 仍取 $\frac{1}{4}(a-c)$)，

获相应的盈利$\frac{9}{64}(a-c)^2$，于是

$$(1-\underline{\delta})\cdot\frac{9}{64}\cdot(a-c)^2+\underline{\delta}\cdot\frac{1}{9}\cdot(a-c)^2=\frac{1}{8}(a-c)^2 \tag{8.41}$$

得$\underline{\delta}=\frac{9}{17}$。

现在关心当$\delta<\frac{9}{17}$的情况。如果两个公司都取"在任何偏离发生之后永远转向 Cournot 产量"这样的触发策略，无疑"适宜的产量应取多少"是个关键问题。首先想到的是这样的触发策略不可能支持一半垄断产量($\frac{q_m}{2}=\frac{1}{4}(a-c)$)那么低。可是众所周知，对任何$\delta$值，简单地一直重复 Cournot 产量策略是子博弈完美均衡。于是我们可以断定，触发策略可以支持的最适宜的产量应当在q_c与$\frac{q_m}{2}$之间，即在$\frac{1}{4}(a-c)$与$\frac{1}{3}(a-c)$之间。设对应于给定的δ($\delta<\frac{9}{17}$)值，q^*是一个适宜的产量策略，考虑如下触发策略：

在第一个周期均取q^*，在任何一个t周期，只要以前未发生局中人偏离现象，则一直取q^*，否则，若发生单方面偏离，则转向 Cournot 产量q_c。

如果两个公司均取q^*，那么阶段博弈(即 Cournot 竞争博弈)中，每一个公司的盈利应为

$$u=(a-c-q_1-q_2)q_1=(a-c-2q^*)q^* \tag{8.42}$$

上述触发策略仍以 Nash 均衡作为惩罚，因此对于"给定的q^*"，由完美无名氏定理的证明，δ应当满足

$$(1-\delta)\max_a g_1(a)+\delta\cdot\frac{1}{9}(a-c)^2<(a-c-2q^*)q^* \tag{8.43}$$

$\max\limits_a g_1(a)$为局中人 1 单方面偏离时可能获得的最大盈利。即求解

$$\max_{q_1}(a-c-q^*-q_1)q_1$$

其解为 $q_1=\frac{1}{2}(a-c-q^*)$，相应的盈利为 $\frac{1}{4}(a-c-q^*)^2$，(8.43)式变为

$$\frac{1}{4}(a-c-q^*)^2(1-\delta)+\frac{1}{9}(a-c)^2\delta<(a-c-2q^*)q^* \tag{8.44}$$

或者说，当 δ 为给定的折扣因子值时，q^* 必须满足(8.44)式才使触发策略为子博弈完美均衡，我们可以试求得 q^* 的适宜范围。经化简，(8.44)式可写作

$$[3q^*-(a-c)][3(9-\delta)q^*-(9-5\delta)(a-c)]<0 \tag{8.45}$$

由于 $q^*<\frac{1}{3}(a-c)$，故

$$q^*>\frac{9-5\delta}{3(9-\delta)}(a-c) \tag{8.46}$$

即 $q^*\in\left(\frac{9-5\delta}{3(9-\delta)}(a-c),\frac{1}{3}(a-c)\right)$。如果令 δ 单调上升，则 q^* 的最低值单调下降。当 δ 上升且接近于 $\frac{9}{17}$，q^* 接近于垄断产量的一半，而当 δ 下降并接近于 0 时，q^* 的最低值明显地接近于 $q_c=\frac{1}{3}(a-c)$，这很容易想象，当 $\delta\to 0$ 时，表明无限重复 Cournot 竞争越来越接近于一次 Cournot 竞争，其合理结局当然应该是静态 Nash 均衡。

上面这段讨论是当 $\delta<\frac{9}{17}$ 时，为使触发策略为子博弈完美均衡，我们求得适宜生产的产量范围。但是在触发策略中，只是以永远转向 Nash 均衡作为惩罚或威胁。前面已指出，人们有理由认为 Nash 威胁太“心慈手软”了些，恐怕不足以阻止局中人违规与偏离。下面仍以无限重复 Cournot 竞争为例，介绍 Abreu 严厉惩罚

思想。固定 $\delta=\frac{1}{2}$（略小于$\frac{9}{17}$），考虑"胡萝卜加大棒"(Carrot-and-stick)策略，以 x 表示最严厉冷酷的惩罚产量：

"在第一周期生产一半垄断产量$\frac{q_m}{2}$。在任意 t 周期，如果在 $(t-1)$ 周期两个公司都生产$\frac{q_m}{2}$，或者都生产 x，那么继续生产$\frac{q_m}{2}$，否则，就生产 x。"

该策略包含两种状态：公司生产 x 的惩罚状态，和公司生产$\frac{q_m}{2}$的合作状态。如果任一公司在某周期偏离合作状态，实际上在该周期内，在对手取$\frac{q_m}{2}$条件下，他生产 $q=\frac{3}{8}(a-c)$ 而获得好处，但是随之惩罚状态开始。如果任一公司偏离惩罚状态，依策略定义，惩罚状态将再一次开始，如果没有一个公司偏离惩罚状态，那么又开始合作状态。

设想两个公司均生产 x，那么每个公司从该周期中获得盈利是 $(a-c-2x)x\doteq\pi(x)$。在下一个周期它们将进入合作状态，如果以后一直合作下去并不发生偏离，那么每个周期内每个公司可指望得到$\frac{\pi_m}{2}=\frac{1}{8}(a-c)^2$。一个子博弈，如果开始周期的结局为 (x, x)，此后一直为$(\frac{q_m}{2},\frac{q_m}{2})$，那么如果以 $V(x)$ 表示每一个公司在该子博弈中获利的现时值（以子博弈开始周期计算），则有

$$V(x)=\pi(x)+\delta\cdot\frac{1}{2}\pi_m+\delta^2\cdot\frac{1}{2}\pi_m+\cdots$$

$$=\pi(x)+\frac{\delta}{1-\delta}\cdot\frac{1}{2}\pi_m \qquad (8.47)$$

如果公司 i 在某周期内将要生产 x，另一公司 j 极大化自己盈利的生产量应当为解

$$\max_{q_j}(a-c-x-q_j)q_j \qquad (8.48)$$

易得解 $q_j=\frac{1}{2}(a-c-x)$，其相应盈利记为 $\pi_{dp}(x)=\frac{1}{4}(a-c-x)^2$，这里 dp 表示“偏离惩罚”。

如果公司 i 在某周期继续生产 $\frac{q_m}{2}$，另一公司 j 企图极大化自己盈利的生产量应当解

$$\max_{q_j}\left(a-c-\frac{q_m}{2}-q_j\right)q_j \tag{8.49}$$

解为 $q_j=\frac{3}{8}(a-c)$，相应盈利记作 $\pi_d=\frac{9}{64}(a-c)^2$，$d$ 表示“偏离合作”。

根据已定义的策略剖面，无限重复 Cournot 博弈中的子博弈可以分为两类：

(1)合作子博弈，它们的前一个周期的结局要么是 $\left(\frac{q_m}{2},\frac{q_m}{2}\right)$，要么是 (x,x)；

(2)惩罚子博弈，它们的前一周期的结周既不是 $\left(\frac{q_m}{2},\frac{q_m}{2}\right)$，也不是 (x,x)。

若要使策略剖面成为子博弈完美，上述两类子博弈中，按照该策略的行动必须是 Nash 均衡。先考虑合作子博弈，欲使局中人不会在第一周期主动偏离策略剖面，相当于每个公司必须宁可一直接受每个周期生产一半垄断产量的盈利，而不愿接受该周期内 π_d(这给它带来短期好处)，却在下一个局期接受惩罚现时价值 $V(x)$，即下式所描述：

$$\frac{1}{1-\delta}\cdot\frac{1}{2}\pi_m\geqslant\pi_d+\delta V(x) \tag{8.50}$$

在惩罚子博弈中，每个公司又必须宁愿执行惩罚而不愿在该周期接受 π_{dp}(它偏离惩罚)并在下一周期再次开始惩罚.用数学公式描述这段话：

$$V(x)\geqslant\pi_{dp}(x)+\delta V(x) \qquad (8.51)$$

将(8.47)式中的 $V(x)$ 代入(8.50)式，得

$$\delta(\frac{1}{2}\pi_m-\pi(x))\geqslant\pi_d-\frac{1}{2}\pi_m \qquad (8.52)$$

(8.52)式的右边表示在偏离的周期所增加的盈利，左边表示在下一周期由于惩罚而招致的盈利损失关于前一(偏离)周期的现时值。(8.52)式表示，若欲使公司不偏离合作状态，必须让违规偏离所得不超过紧接着的下一个惩罚周期蒙受损失的现时价值(假定没有一个公司偏离惩罚状态，于是自下一周期以后不存在损失，因为惩罚结束，两公司回到垄断结局，仿佛没有过偏离那样)。类似地，(8.47)式中的 $V(x)$ 代入(8.51)式，可得

$$\delta\left(\frac{1}{2}\pi_m-\pi(x)\right)\geqslant\pi_{dp}-\pi(x) \qquad (8.53)$$

$\pi_{dp}-\pi(x)$ 表示公司从偏离惩罚状态所增加的盈利，它不应当超过"如果这一周期不偏离惩罚状态而在下一周期内可以增加的盈利的现时价值"($\left(\frac{1}{2}\pi_m-\pi(x)\right)$ 也可解释为，由于公司偏离惩罚状态，而在下一周期重新进入惩罚状态所招致的损失)。(8.52)式与(8.53)式这两个不等式的解释完全符合人们正统的思维方式。

当 $\delta=\frac{1}{2}$ 时，(8.50)式应为 $\pi_m\geqslant\pi_d+\frac{1}{2}V(x)$，以 $\pi_m=\frac{1}{4}(a-c)^2$，$\pi_d=\frac{9}{64}(a-c)^2$，$V(x)=(a-c-2x)x+\frac{1}{8}(a-c)^2$ 代入此式，得

$$[8x-(a-c)][8x-3(a-c)]\geqslant 0 \qquad (8.54)$$

即得 $x\geqslant\frac{3}{8}(a-c)$ 与 $x\leqslant(\frac{1}{8})(a-c)$

再考虑(8.51)式，当 $\delta=\frac{1}{2}$，$\pi_{dp}=\frac{1}{4}(a-c-x)^2$，因此(8.51)式为

$$\frac{1}{2}V(x)\geqslant\pi_{dp}(x)$$

即

$$(a-c-2x)x+\frac{1}{8}(a-c)^2 \geqslant \frac{1}{2}(a-c-x)^2$$

经整理得

$$[2x-(a-c)][10x-3(a-c)]\leqslant 0 \quad (8.55)$$

可得

$$\frac{3}{10}(a-c)\leqslant x\leqslant \frac{1}{2}(a-c)$$

将上述 x 的可能范围合并，得

$$\frac{3}{8}(a-c)\leqslant x\leqslant \frac{1}{2}(a-c) \quad (8.56)$$

当 x 属于(8.56)式范畴，策略剖面为子博弈完美均衡。$x\leqslant\frac{1}{2}(a-c)$是显然且必然的事情。请注意 $x\geqslant\frac{3}{8}(a-c)$，回顾(8.49)式，当公司 j 偏离合作状态时，它的最佳短期盈利的行动选择为 $q_j=\frac{3}{8}(a-c)$，由此它将招致的惩罚是 $x>\frac{3}{8}(a-c)$的产量。简单的道理：你想多生产一些，那么招致大家生产得更多，从而受到的惩罚越重。

3. 效率工资(efficiency wages)

我们在前面介绍的无限重复博弈，基本特征是阶段博弈为静态博弈，但我们也曾提及，当阶段博弈为动态时，讨论的结果仍然成立。本小段讨论的效率工资模型，其阶段博弈就是动态的。

在效率工资模型中，公司劳动力的产品依赖于公司支付的工资，在发展中国家，较高的工资可导致较好的营养；而在发达国家，较高工资可以吸引能力强的人员来公司工作，或促使在职工人努力工作。

Shapiro 与 Stiglitz 于 1984 年提出了一个动态的工资模型，公

司通过高工资促使工人努力工作，并且威胁开除开小差者。由于采用高工资策略，公司削减对劳动力的需求，于是有些人被公司以高工资雇用，而另一些人则(不情愿地)遭到解雇。失业工人的队伍不断庞大，失业工人寻求新工作所花费的时间也越来越长，于是解雇威胁变得更有效。在竞争(竞赛)均衡中，工资 w 与失业率 u 导致工人不偷懒，确定工资水平 w 的公司劳务需求导致了失业率恰好为 u。这里讨论的无限重复博弈中的阶段效率工资模型只分析一个公司与一个工人的情况。

考虑下述阶段博弈：

首先，公司向工人开价工资 w。其次，由工人接受或拒绝公司的出价。如果工人拒绝 w，则他以工资 w_0 自我雇用成为个体户。如果工人接受 w，那么他选择或者努力工作(此时需承担不利因素 e，譬如人太累需要补充营养，或者其他额外的支出)或者偷懒(自然不必承担不利因素)。公司可能观察不到工人是否努力工作的决策，但是工人的产量将被公司和工人双方都观察到，假定产量有高低之分，为简便起见，采用"标准化"形式，将低产量取为0，因此，高产量价值 y 必定是大于0的正数，一个合理的假设为：如果工人努力工作则肯定会有高产量；如果工人偷懒，那么以概率 p 为高产量而以概率 $(1-p)$ 产量为0，毫无疑问地，低产量是工人偷懒的一个信号。

现在确定公司与工人的盈利函数。设公司以工资 w 雇用工人，如果工人努力工作并获得高产量，那么公司得 $y-w$，而工人得 $w-e$；如果工人偷懒，e 自然为0；倘若呈现低产量，则 y 也自然为0。在这种场合下，工人的盈利显然是 w，而公司的期望盈利则为 $py-w$。如果工人被公司雇用并努力工作是有成效的话，至少 $y>w$，且 $w-e\geqslant w_0$，故此时有 $y-e>w_0$。又若工人自谋出路比起被公司雇用但偷懒的处境要好一些的话，当有 $w_0>py$，因此在模型中，我们假定 $y-e>w_0>py$。

该阶段博弈既然是动态的，我们自然首先关心它的子博弈完美均衡。在这个动态模型中，公司首先行动，由公司预付工资 w 给工人，此时工人采取（偷懒，努力工作）两个策略之一，其盈利分别是 w 与 $w-e$，因此公司预测工人必然没有激情努力工作，于是公司将预付 $w=0$（或任意小于 w_0 的 w），而对此工人的反应是自谋出路。尤其是在 p 相当小，$py-w$ 可能为负数时，该策略剖面是子博弈完美——一个相当惨淡凄凉的结局。但是，如果阶段博弈无限重复地进行，情况与一次博弈有极大差别。公司可以通过支付超过 w_0 的 w，并且如果产量一直低的话，则以开除作为威胁来促使工人努力工作。我们将证明，对某些参数值，公司发现通过支付奖金来促使工人努力工作是值得的一举。

考虑无限重复博弈的下述策略，它包含了以后待定的工资 $w^* > w_0$。对于历史 h^t，只要所有以前的开价均为 w^*，并都被工人接受，而且每一周期工人都努力工作使产量很高，则称该历史为高工资、高产量的。公司的策略是，在第一周期开价 w^*，在以后的各个周期，如果历史是高工资、高产量的，则继续开价 w^*，否则开价 $w=0$；工人的策略是，如果 $w \geq w_0$ 则接受公司开价，否则就选择另行自谋出路；如果历史是高工资、高产量的，且当前开价为高工资，则选择努力工作，否则就偷懒。

公司的策略类似于触发策略：假如以前一切行动都取合作则继续合作，若合作一旦遭受破坏，则转向阶段博弈的子博弈完美结局（即取 $w=0$）。工人的策略也类似触发策略，但我们必须注意到阶段博弈是动态时与静态情况有所不同。因为工人在动态博弈中处于第二位行动者，因此作为第一位行动者的公司，如果偏离的话则在阶段博弈进程中就可被察觉，而不像静态阶段博弈中的可能偏离现象必须在阶段博弈结束时才察觉。因此一旦公司偏离，工人将在该周期就作出相应反应，绝不会等到下一周期。假如以前所有的行动均取合作态度，那么工人的策略是继续合作，可是对于公司

的偏离，工人将作出相应反应，并知道阶段博弈的子博弈完美结局将在所有的以后阶段里实施。特别地，如果 $w \neq w^*$，但是 $w \geqslant w_0$，那么工人接受公司的开价但是偷懒。

那么，在什么样的条件下这些策略是子博弈完美 Nash 均衡呢？我们先来看看策略为 Nash 均衡的条件，所使用的方法是从 Nash 均衡的定义出发，即在我们所讨论的策略剖面中，固定某一个局中人策略时，求另一个局中人的策略的确是最佳反应所需要的条件。

假定公司在开始的第一周期开价为 w^*，在给定公司这一策略下，依策略剖面的定义，只要 $w^* \geqslant w_0$，工人接受 w^* 是最佳行动，在该前提下工人有两种选择：如果工人努力工作，势必呈现高产量，于是公司继续提供 w^*，并且工人将面对下一局期同样地提供努力决策。倘若努力工作对工人来说是最佳反应，那么他所得到的盈利的现时价值为

$$V_e = (w^* - e) + \delta V_e \tag{8.57}$$

下标 e 表示努力工作(effort)。(8.57)式也可写成

$$V_e = (w^* - e)/(1 - \delta) \tag{8.58}$$

倘若工人在接受 w^* 以后采取偷懒策略，他将以概率 p 生产高产量，此时将在下一个周期提出同样的提供努力的决策，可是工人又以概率$(1-p)$生产低产量，公司在观察到低产量以后将从下一局期开始永远开价 $w=0$，这样工人也就永远地自谋出路。据上述分析，如果工人认为偷懒是最佳方案，那么工人盈利的（期望）现时价值为

$$V_s = w^* + \delta\left\{PV_s + (1-p)\frac{w_0}{1-\delta}\right\} \tag{8.59}$$

这里下标 s 表示偷懒或开小差(shirk)，V_s 也可写成

$$V_s = [(1-\delta)w^* + \delta(1-p)w_0]/(1-\delta p)(1-\delta) \tag{8.60}$$

如果 $V_e \geqslant V_s$ 的话，那么工人努力工作是最佳策略，即

$$w^* \geqslant w_0 + \left[\frac{1-p\delta}{\delta(1-p)}\right]e = w_0 + \left[1 + \frac{1-\delta}{\delta(1-p)}\right]e$$

$$= w_0 + e + \left[\frac{1-\delta}{\delta(1-p)}\right]e \tag{8.61}$$

(8.61)式表示，为了引导工人努力工作，公司不仅应当付给工人 w_0+e——自谋出路的收入加上努力工作所花费的成本，而且还要支付给工人额外奖金 $\frac{(1-\delta)e}{\delta(1-p)}$。当 p 非常接近于 1 时，表示公司很难发现工人偷懒，因此为引导工人努力工作，公司必须付出极其高的奖金。如果 $p=0$，我们不能使用公式(8.61)式，因为几乎必然发生的低产量使得公司将永远辞退工人，让他自谋出路。因此在 $p=0$(极易被发现偷懒)时，如果

$$\frac{1}{1-\delta}(w^* - e) \geqslant w^* + \frac{\delta}{1-\delta}w_0 \tag{8.62}$$

则工人宁愿不偷懒而选择努力工作。(8.62)式的左边为无限重复效率工资博弈中每周期均努力工作的工人的盈利函数，右边则为工人在第一周期偷懒而此后只能当个体户的盈利。(8.62)式可以写成

$$w^* \geqslant w_0 + \frac{e}{\delta} = w_0 + (1 + \frac{1-\delta}{\delta})e \tag{8.63}$$

非常有趣的是，它恰好与(8.61)式中取 $p=0$ 时完全一样。

即使(8.61)式成立，使得工人的策略是关于公司工资策略的最佳反应，但对公司来说，支付工资 w^* 也一定是值得的。

由于博弈存在两个局中人，因此我们还需要讨论给定工人策略的情况。在第一周期，公司面临两种选择：

(1)支付 $w=w^*$，并且当低产量在任何时候被观察到时，以这样的高工资 w^* 威胁开除工人(注：工资 w 低时，“开除”构成不了威胁)，从而引导工人去努力地工作，由此在每一个周期，公司均可获得 $y-w^*$。

(2)支付 $w=0$，由此导致工人自谋出路，从而在每周期接受盈利为 0。

显然，如果

$$y-w^{*}\geqslant 0 \tag{8.64}$$

公司的策略是关于给定工人策略的最优反应。我们在开头已经假定了 $y-e\geqslant w_0$，这只是效率工资模型的基本且合理的假设。现在我们将看到，若要所叙述的策略剖面是 Nash 均衡，(8.61)式与(8.64)式至少得同时成立，于是综合这两个不等式不难得到

$$y-e\geqslant w_0+\left[\frac{1-\delta}{\delta(1-p)}\right]e \tag{8.65}$$

(8.65)式显然比 $y-e>w_0$ 更具体了一些，或者说(8.65)式蕴含了 $y-e>w_0$，但反之未必如此。(8.65)式可以解释为我们已经熟悉的限制：如果合作将持续，那么 δ 必须充分地大。因为 δ 越趋近于 1，(8.65)式无疑地越来越趋于模型原来的基本假设 $y-e>w_0$，也就是说，只要 δ 充分地大，在效率工资模型的基本假设下，我们构造的策略剖面才有可能成为子博弈完美均衡。

我们已经从 Nash 的定义出发，证明了只要(8.61)式与(8.64)式这两个不等式成立，所确定的策略剖面至少是 Nash 均衡。接下来就要关心这些策略是否为子博弈完美。首先我们观察一下这里的重复博弈的子博弈的特点。回想前面介绍的以静态博弈(即局中人同时行动)作为阶段博弈的无限重复博弈，其子博弈开始于两个周期(或阶段)之间，而我们这里考虑的无限重复博弈，其阶段博弈是动态的，局中人的行动是相继的，有先后顺序的。因此可以设想子博弈非但可以从两个周期之间开始，也可以从阶段博弈的内部开始，例如，从某一周期工人观察到公司的开价以后再作决策开始的子博弈就是从阶段中间开始的。在我们构造的策略剖面中，可以将子博弈分为两类：

(1)在高工资高产量的历史之后开始的子博弈；

(2)在所有其他历史之后开始的子博弈。

我们在前面证明了确定的策略剖面是 Nash 均衡，这个证明过程实际上也证明了确定的策略剖面在第(1)类子博弈中是 Nash 均衡。如果我们也能证明所确定的策略剖面在第(2)类子博弈中也构成 Nash 均衡，事实上我们已经从基本定义出发证明了策略剖面是子博弈完美均衡。

在第(2)类子博弈中，由于历史不是高工资、高产量的，因此工人将绝不会卖力地干，公司最好的办法是导致工人选择自谋出路。从而在下一阶段乃至永远，公司将开价 $w=0$，工人在本阶段中应当不努力工作并仅当 $w \geqslant w_0$ 接受眼前的开价。根据这两方面的分析，在历史不为高工资、高产量时，给定工人的策略(不努力工作)，公司最佳策略为开价 $w=0$，这正是策略剖面中所规定的策略。于是，策略剖面在第(2)类子博弈中构成 Nash 均衡。

由于我们假定阶段博弈中只有一个公司与一个工人，因此上面的例子作为无限重复博弈的一类讨论是合理的，而所得结论却与现实相距甚远。在我们的子博弈完美均衡中，自谋出路是永久性的：如果工人在任何时候被逮到偷懒，公司从此以后开价 $w=0$；如果公司在任何时候偏离开价 $w=w^*$，那么工人将不再努力工作，公司也就不再承担雇用工人的费用。这种永久性自谋出路是否合理呢？由于公司与工人的单一性，显然双方宁愿返回无限重复博弈中的高工资、高产量均衡，因为这对双方都有好处，他们均不会愿意阶段博弈中的子博弈完美结局—— 在那里，公司开除工人而失去了劳动力，工人失去了工作而只能自谋出路当个体户，这对双方都没有什么好处。这里面存在着一个再谈判的问题(有关再谈判内容，我们将另立一节详细进行讨论)。其实，在单一的阶段博弈中，由于像重复博弈中的惩罚不会付诸实施，因此，类似重复博弈中由此类惩罚的威胁所导致的合作不再是均衡。

在正式的实际劳动力市场，公司雇用许多工人，对于一个工人

的上述情况，公司是不愿意重新谈判的，因为与一个工人的重新谈判可能搅乱仍与其他工人进行的(或已经开始的)高工资、高产量均衡。假如有许多家公司，问题又转成公司 j 是否录用被公司 i 正式雇用的工人。也许公司 j 不会，因为它担心这样做会搅乱它与自己目前雇用工人之间的均衡，恰如在单个公司情况一样。

此外，如果解雇的工人总是可以找到他们乐于自谋出路的新工作，那么，正是这些新工作的工资在这里起着自谋出路中工资 w_0 的作用。在被解雇工人根本未遭受损失的极端情况下，就不存在在无限重复博弈中对可能的偷懒实施惩罚，因而也不存在工人努力工作的子博弈完美均衡。1989 年 Bulow 与 Rogoff 在债权问题上的类似思想有一个精彩的应用：如果一个负债国可以通过国际资本市场中短期预付现金交易接受它从债权国家得到的长期贷款，那么，对于在负债国与债权国之间的无限重复博弈中的可能违约行为不存在惩罚。

§8.4 对手变化的重复博弈

前面介绍的重复博弈中，阶段博弈完全相同，其中局中人也完全相同。但在日常社会经济活动中，每个周期中局中人不是固定不变的。例如，某公司生产某产品，每个周期公司与顾客有买卖之间的交易，但在各个周期，顾客可能有所不同，有一些是长期顾客，每个周期都来交易，而另一些角色则由一些短期或临时顾客扮演。我们关心的是，在对手发生变化的重复博弈中，无名氏定理的结论是否仍然成立。本节将分三种情况进行讨论。

1. 含有长期与短期局中人的重复博弈

我们考虑的是无限重复博弈，因此需要如前面一样地建立模型。在前面的模型里，长期局中人为标准重复博弈中的周中人，他们周期复周期地参与博弈。另外的短期局中人序列，他们中的每一

个人只进行一个周期的博弈。为了对问题的展开有直观了解，不妨以两个例子加以说明。

例 8.2 长、短期局中人的最实际例子是公司（或厂家）与顾客。简单地考虑，一家公司面对一系列顾客，公司是长期买卖博弈中的长期局中人，顾客是短期局中人。假设一个顾客只买一次公司的产品——即只参加一次交易博弈，但是每一个顾客在选择他自己的行动（譬如，买或不买）时，得到以前所有交易的有关知识。假如在每一个周期，总是顾客首先行动，由他选择是否购买公司的产品。如果顾客不买，那么两方局中人获益为 0。如果顾客决定购买，那么由公司决定是生产高质量产品还是低质量产品。如果生产高质量产品，公司与顾客将获盈利 1；如果公司生产低质量，则公司盈利为 2，而顾客盈利为 −1。倘若只考虑一次性交易，显然这是一个动态博弈，顾客预测，他如果决定购买，公司必然生产低质量产品，于是顾客决定不购买。但是无限次重复博弈将会有怎样的结果呢？注意到公司是长期局中人，需要考虑折扣因子 δ，但是顾客是短期局中人，每个顾客仅在自己参与交易的局期中极大化自己的盈利。基于这些思路，考虑下述策略：

“公司从顾客购买的每一次开始生产高质量产品，只要在以前绝无生产过低质量产品，就继续生产高质量产品。如果，在任何时候公司生产低质量产品，那么它在以后每一个后继机会中均生产低质量产品。

第一个顾客购买公司的产品，只要公司没有生产过低质量产品，后续的顾客在以后的周期都选择购买。如果在任何时候公司生产低质量产品，以后没有一个顾客会去购买公司的产品。”

我们将说明，只要公司有充分的耐心（即 δ 比较大），那么上述策略是子博弈完美均衡。注意到这些策略与“阶段博弈”的预测结局存在着显著的差异，这里的差异类似于前面关于无限重复博弈的讨论。

在给定公司策略下，顾客的策略是最佳反应。因为每一个顾客只关心他所参与交易周期的盈利，因此，当且仅当该周期的质量如期望的那样高，顾客会购买公司的产品。再考虑给定顾客策略时的情况，公司在只与一个顾客成交(即顾客先表示购物)时，如果生产高质量产品将比生产低质量产品损失 1(即少获盈利 1)，但是生产低质量产品将赶走将来的顾客。设想从第 1 周期至第 k 周期，公司生产高质量产品，而在第$(k+1)$周期生产低质量产品，再往后的周期没有顾客前来购货，此时公司的折扣盈利和应为

$$1+\delta+\cdots+\delta^{k-1}+2\delta^{k}=\frac{1-\delta^{k}}{1-\delta}+2\delta^{k} \tag{8.66}$$

若要(8.66)式大于 2(在第一次交易时公司就生产低质量产品)，仅需 $\delta>\frac{1}{2}$即可。且在 $\delta>\frac{1}{2}$时，(8.66)式总是小于永远生产高质量的盈利的折扣和。因此，当公司具有一定的耐心时，在给定顾客的策略下，上述公司的策略的确是最佳反应。于是我们至少证明了所述策略是 Nash 均衡。

剩下的问题仍要证明这些策略是子博弈完美的，必须注意到形式上的“阶段博弈”是动态的。因此我们可以将无限重复博弈的子博弈分为两类：历史为(购买，高质量)的子博弈，策略剖面在这类子博弈中的相应部分，仅需利用上述分析就可知是 Nash 均衡。另一类为历史呈现不购买与非高质量子博弈，策略剖面在这一类子博弈中为 Nash 均衡是显然的，读者不妨自行做个小练习。

例 8.1 很能体现日常生活中的买卖现象。如果阶段博弈是动态的，由顾客先提出是否购买，这相当于顾客向一家固定的商店提出购买要求。该情况的子博弈完美均衡指出：商店(或公司)为了长期的利益，比较注重信誉，因此它每次将以高质量的产品出售给顾客。又如果阶段博弈是静态的，那么卖方实际上相当于串街走巷的小贩，每次买卖的可能结局是：顾客不购买，小贩提供低质量产品。子博弈完美均衡指出，一般顾客不喜欢到这些小摊贩上购物。

例 8.3 考虑囚徒窘境中局中人的行动分先后次序的情况。一个长期局中人面对一系列短期局中人。在每一周期博弈中，短期局中人关于是合作还是欺骗背叛作出决策，长期局中人观察到上述决策行动之后，再作出自己的决策。经验告诉我们，如果长期局中人是个"老兵油子"，那么无限重复博弈的子博弈完美均衡是：囚徒总是互相合作——即拒不交代。这个均衡用策略来完整地描述："短期局中人取合作策略，只要在每一个过去的周期内，长期局中人采取与那个周期的短期局中人相同的行动；如果在任何时候长期局中人的行动与当时的短期局中人行动不一致，则短期局中人取欺诈背叛策略。对于长期局中人来说，只要他在过去的周期从未与当时的对手行动不一致，那么他将采取与当前周期对手相同的行动，否则他就采取欺诈背叛策略。"

这两个例子有若干共同特点：一个长期局中人面对一系列短期局中人；每个短期局中人只在一个周期参与博弈；每个周期(即阶段博弈)局中人的行动不是同时发生的，短期局中人先行动，在观察到短期局中人的这一行动以及知道以前各个周期大家的行动后，长期局中人作出自己的决策。如果在阶段博弈中局中人无论长期还是短期，都同时行动，情况很可能有所不同。譬如在例 8.2 中的合作均衡中，由于短期局中人首先行动，因此不必考虑他们采取的合作态度是由今后周期的奖罚而激励出来的，他们只博弈一次，无论吃亏还是便宜，下一次不再光临，当然在每个周期的博弈中，他们总是希望极大化自己的盈利。而长期局中人则必须考虑到今后周期的效用，信誉至上对于他的长期利益至关重要，只要对手不欺诈，他的最佳反应也是合作。出于这一点，短期局中人采取合作均衡策略。如果在阶段博弈中，双方行动将同时选择，尽管阶段博弈无限次重复进行，但短期局中人只博弈一次，在一次静态囚徒窘境形博弈中，"大家总是采取欺诈背叛"是唯一的 Nash 均衡，因此

“博弈双方总是欺诈背叛”构成无限重复博弈中唯一的子博弈完美均衡。这个结论与无限重复囚徒窘境博弈的无名氏定理结论发生了根本性的变化。因为局中人有长、短期之分，短期局中人着眼于短期效用，因此从原先讨论过的可行盈利集合出发考虑问题的做法不适合具有长、短期局中人的重复博弈，要使无名氏定理可以推广到这一类重复博弈，思路之一就是修改可行盈利的定义，随之也就要修改与之相关的 minmax 值定义。修改的目的是使得“短期局中人总是采取最佳反应”的约束达到具体化，落实到模型上。阶段博弈的局中人将分为两部分：长期局中人 $i=1,2,\cdots,\lambda$，与短期局中人 $i=(\lambda+1),\cdots,n$。长期局中人的目的在于使他们的每周期盈利的标准折扣和达到极大，短期局中人则只关心他所行动的周期盈利达到极大。综上所述，每次阶段博弈有 n 个局中人，但是任何两个阶段博弈中的 $(n-\lambda)$ 个短期局中人是完全不相同的，因为我们的模型规定每个短期局中人只参与一次阶段博弈，这是真正意义的“短期”！为了落实“短期局中人总是采用短期最佳反应”这段话，我们不考虑 n 个局中人所有可能结局的盈利向量，合理的办法是对 λ 个长期局中人，考虑其所有可能的行动策略，而对短期局中人，只考虑在给定其他局中人的策略下，局中人 $i(\geqslant\lambda+1)$ 的最佳策略——Nash 均衡策略。基于此点，我们构造如下图形：设

$$B:\mathscr{A}_1 X\mathscr{A}_2 X\cdots X\mathscr{A}_\lambda\rightarrow\mathscr{A}_{\lambda+1}X\cdots X\mathscr{A}_n$$

为将长期局中人的任意行动策略剖面 $(\alpha_1,\cdots,\alpha_\lambda)$ 映照到与该剖面响应的短期局中人的 Nash 均衡行动的一个对应。即，对每一个 $\alpha\in$ 图(B) 和 $i\geqslant\lambda+1$，α_i 是关于 α_{-i} 的最佳反应。现在我们要定义的 minmax 值当然建立在图(B)的基础上，注意到短期局中人的目光仅盯住每周期阶段博弈的 Nash 均衡盈利，因此修正 minmax 值 $\underline{v}_i$ 仅对长期局中人 $i(=1,\cdots,\lambda)$ 定义：

$$\underline{v}_i=\min_{\alpha\in\text{图}(B)}\max_{a_i\in A_i} g_i(a_i,\alpha_{-i})\qquad i=1,2,\cdots\lambda\qquad(8.67)$$

由于图(B)是紧集且盈利函数在混合策略空间是连续函数，因此(8.67)式最后取极小值是可达到的。读者不妨验证，如果所有n个局中人都是长期的，那么(8.67)式中的定义回到了原先无限重复博弈的 minmax 值的定义。

接下来，我们关心修正可行盈利集的定义。令

$$U=\{v=(v_1,\cdots v_\lambda)\in R^\lambda \mid \exists\ \alpha\in 图(B), 且\ g_i(\alpha)=v_i, i=1,\cdots,\lambda\} \tag{8.68}$$

注意这样一个事实，U尽管是λ个长期局中人的盈利向量集合，但由于条件$\alpha\in$图(B)的限制，在所讨论的问题中，短期局中人的短期最佳行为的限制已经包含在其中。置

$$V=U 的凸包 \tag{8.69}$$

这就是可行盈利集的修正定义。

有了修正的可行盈利集与 minmax 值的修正定义，按理我们可以设法推广无名氏定理到具有长短期局中人的无限重复博弈中去。然而，Fudenberg、Kreps 与 Maskin 于 1990 年证明了如下事实：仅当阶段博弈中每一个局中人的混合策略的选择公开地可观察到，方可得到无名氏定理的推广。当局中人仅仅观察到他们的对手已实施了的行动，子博弈完美均衡集的范围可能严格地较小。

假定

$$\overline{v}_i=\max_{\alpha\in 图(B)}\ \min_{a_i\in(\alpha_i)的支撑} g_i(a_i,\alpha_{-i}) \tag{8.70}$$

对于一个固定的$\alpha\in$图(B)，相应的α_i的支撑点是指混合策略(即概率分布)取正概率的几个纯策略，(8.70)式后半部分求极小值表示在所有$g_i(a_i,\alpha_{-i})$(a_i为α_i赋予正概率的纯策略)求最小的数值，这个数值实际上是局中人i在给定$\alpha\in$图(B)时的最差的盈利(他当然小于局中人i在α下的期望盈利)。每一个$\alpha\in$图(B)，局中人i存在相应的最差盈利，$\overline{v}_i$无非是在所有可能的最坏结果中挑选一个α使局中人i“倒霉”程度为最小的盈利。

推广的无名氏定理中的诸 v_i 非但与$\underline{v_i}$ 有关，且与$\overline{v_i}$有关，我们不加证明地介绍有关结果：

定理 8.11 (Fudenberg, Kreps, Maskin 1990)假定 V 的维数等于长期局中人的个数，那么对每一个 $v\in V$ 且$\underline{v_i}<v_i<\overline{v_i}$，$i=1,2,\cdots,\lambda$，存在一个$\underline{\delta}$ 使得对所有 $\delta\in(\underline{\delta},1)$，存在一个子博弈完美均衡，其盈利为 V。

2. 叠代博弈(overlapping generations)

考虑这样的重复博弈，每个周期有 T 个"时期"的局中人一起博弈，在每个时间 t(或第 t 个周期)，有一个"年龄"为 T 的周中人正在进行他的最后一轮博弈；"年龄"为$(T-1)$的局中人除了这一轮进行博弈外，还有一轮有待博弈；一直推断下去，直到一个新的局中人，其"年龄"为 1，在 t 周期刚开始他的第 1 次博弈，并且他还有$(T-1)$次周期将要进行博弈。每个周期，T 个不同年代的局中人一起博弈，所以称为叠代博弈，他们同时选择策略行动：工作或偷懒。他们的选择在周期结束时展现，局中人均分他们在该周期的成果，成果收益是选择参加工作的局中人数的递增函数。至于努力工作的成本花费超过享有所增加收益的 $1/T$ 倍。显然，偷懒是阶段博弈中的占优策略，这有点像 T 个周中人的囚徒窘境的味道。重复博弈中每个局中人的盈利定义为周期效用的平均数。

假定有效结局是所有局中人工作，可是这个有效结局不可能发生在任何 Nash 均衡中，因为无论在哪一个周期，年龄为 T 的局中人将总是偷懒，然而，可能存在大多数局中人工作的均衡。例如，令 $T=10$，假如 k 个局中人工作则总计产量(收益)为 $2k$，以及努力的负效用为 1，于是，当有 k 个对手工作时，局中人工作的盈利应为 $2(k+1)/10-1$，如果他偷懒则得盈利 $2k/10$。有效结局是所有 10 个局中人都工作，每个局中人获盈利 1。

考虑下述策略剖面："'10'岁的周中人总是偷懒。只要没有一

个年龄小于 10 的局中人曾经偷过懒，所有年龄小于 10 的局中人均工作。如果某个年龄小于 10 的局中人曾经偷懒，那么所有局中人都偷懒。”

如果所有局中人遵照这个策略剖面，那么某局中人在他工作的周期获益 18/10－1＝4/5(他之所以工作，是因为其他年龄小于 10 的局中人都工作)，而当他年龄为 10 时的周期他必定偷懒，若其他局中人工作，则获益 9/5——这是 10 岁局中人的最大可能获益。因此，任何一个局中人，当他年龄为 10 时，即使他不偷懒而工作，最多只能获盈利 2×10/10－1＝1＜ 9/5，显然，此时不会因为偏离了‘偷懒’策略而获得好处。如果年龄为 9 的局中人偏离策略剖面而偷懒，那么在他偏离的那个周期他获得 2×8/10＝8/5，但到了下一个周期，由于他在这一周期的偏离而造成大家偷懒，其盈利为 0，但如果他不偷懒的话，两个周期的盈利总计(4/5＋9/5)肯定大于(8/5＋0)。通过计算不难明白，年龄越轻，局中人因偏离策略剖面而造成的损失也就越大。上述分析对任何一个时间 t 开始的子博弈都是一样，故我们证明了所提出的策略剖面是子博弈完美均衡。

这个均衡要求人们，年轻时应当多干些，到了老年可以理所当然地享享福——一种受到绝大部分人认可的良好社会风尚。

3. 随机匹配对手(randomly matched opponents)重复博弈

现在我们讨论这样的无限重复博弈，有两个总体，各有 N 个局中人。每一个周期，从总体 A 中随机地选取一个作为局中人 1，又同时从总体 B 中随机地选取局中人 2，这一对随机匹配的对手参加该周期的二人阶段博弈。每次从总体 A(或 B)的抽样属于随机有放回的，因此，在无限重复博弈中，每个局中人可能无限经常地出现在阶段博弈中。

由于重复博弈的特点是，阶段博弈开始之前已经观察到以前各阶段(或周期)的行动与结局。因此，在随机匹配的阶段博弈中，

每个局中人将非常关心对手是否参加过前面的博弈，在这些参加过的博弈中，该对手的行为与习惯如何，一般地，尤其对他最近一次的行为感兴趣，因为有可能认为最近的博弈行为对当前博弈最有影响。而这一点恰好可以用概率论中的 Markov 性质给予描述。

上述讲法与其他一些合理的想法都可以成为建立随机匹配模型的基本假设，不同的假设构成了不同类型的随机匹配模型，在 Rosenthal 等人的模型中，他们假设了，匹配每一对局中人时，这对局中人的信息由双方在前面周期中的两个局中人采取的策略行动所组成。如果阶段博弈是囚徒窘境，我们以 C 表示“合作”，以 D 表示“背叛”，那么一对局中人有四个可能“历史”，(C,C)，(D,C)，(C,D) 与 (D,D)。倘若局中人没有完美记忆，则每个局中人将具有 16 个纯策略。

根据这样的信息结构，可以有各种可行策略，譬如，“当且仅当对手在上述周期合作则合作”或者“针锋相对”。更一般地，局中人在 t 周期选择的行动将直接影响 $(t+1)$ 周期中他的对手的行为。

如果局中人预测将在 $(t+1)$ 和 $(t+2)$ 周期面对相同的对手，他可能预测他的 t 周期行动将对他的对手在 $(t+1)$ 周期以后的行动有额外的非直接影响。例如，如果对手仅当历史为 (C,C) 时取合作，那么局中人在 t 局期的背叛不只是仅使对手在 $(t+1)$ 周期背叛，而且也使得对手在今后周期都采取背叛行动，这种非直接影响使问题的研究复杂了一些。1979 年，Rosenthal 以及 Rosenthal 与 Landau 将问题的讨论局限于“Markov 均衡”，在这类均衡中不再出现上述非直接影响，每一个局中人相信，他在 t 周期的行动对于他的对手在 $(t+2)$ 周期及以后的所有周期的行动没有影响。请注意这里的所谓 Markov 性质与我们通常在概率论中遇到的 Markov 性在提法上的差异，两者不可完全混为一谈。概率论中的 Markov 性是指 t 时刻随机事件的发生仅与前一时刻的结果有关。而这里是指乙在 t 时刻的行动仅受到甲在前一时刻行动的影响。

倘若两个局中人总体均具有无穷无尽甚至认为集合具有连续统，那么上述 Markov 性的信念必定正确，因为局中人遇到同一个对手的次数在 2 次以上的概率为 0。不过，在 Rosenthal 等人的研究中局中人总体是有限的，因此某局中人 1 与同一个局中人 2 在连接两个周期中匹配的概率不等于 0。简单地设想一下，假设两个总体各只有 2 个人，显然，局中人 1 在 t 时期的行动对于下一个周期他的对手(不管是哪一个)总有一些影响，但是如果总体的人越来越多。那么这种影响也就越来越小。不过，在一个极端的情况下，两总体各只有一个局中人，Markov 性的信念显然不正确。相信读者对此很容易理解。

现在考虑阶段博弈为囚徒窘境的随机匹配重复博弈，假定所有的局中人采取"针锋相对"策略。那么在什么条件下，这个策略剖面是 Markov 均衡？

所谓针锋相对，意即每一个局中人在他的当前对手在上周期采取合作的情况下，必定愿意取合作策略；如果对手在上周期背叛，那么他一定也背叛。然而，Markov 性使得下一个周期的对手不知道这周期之前乙方局中人的所作所为，例如，局中人甲在本周期背叛，下一个对手无法区分到底是由于乙方局中人在再前一阶段背叛而引起甲的背叛，还是甲偏离了"针锋相对"策略而采取背叛。不管怎样，遵循针锋相对策略，下一个对手在看到甲背叛后必定背叛，今天的任何背叛将会导致下一个对手背叛。因此，双方均使用"针锋相对"策略的话，如果某一个局中人背叛，他因此而获益 a，必然招致下一次的惩罚——下一个对手也背叛，从而造成损失 b，其贴现损失为 $b\delta$。要使"针锋相对"策略成为 Markov 均衡，仅当 $b\delta=a$ 才有可能。为什么呢？因为当 δ 较小时，$b\delta<a$，这表示对背叛的惩罚力度太少，以至于不能强制他合作从而可能发生偏离，因此 δ 太小是不行的；那么 δ 相当大总行了吧？当 δ 较大时，惩罚力度相当大，使惩罚者相当担心自己将来的盈利也会减少而不愿惩

罚，从而又偏离了“针锋相对”策略，只有 δ 满足 $a=b\delta$ 时，上面两种偏离可能才不会发生。因此 $\delta=a/b$ 是针锋相对成为 Markov 均衡的条件。

读者可以从囚徒窘境的盈利矩阵直接计算相应的 δ 值。那么 $\delta\neq a/b$ 就一定不存在针锋相对式的 Markov 均衡了吗？倘若在囚徒窘境中双方一直采取背叛，这是一个均衡，此时前面分析的两种可能将会自然消失，谁也不会主动偏离这个均衡，而这个均衡的确是针锋相对式的。

在随机匹配博弈中，如果每一个局中人可以观察到以前阶段博弈的行动，再假定折扣因子 δ 接近于 1，那么合作是均衡结局(Kandori 1989)。对于每一个局中人，他应采取的策略是：

“在第一周期合作，只要我以前每一次参与的博弈结局是(c，c)，以及我现在的对手上一次的博弈结局是(c，c)，那么我在本周期内继续合作；否则我就背叛。”这个策略的基本逻辑是，只要以往我没有骗过人同时也没有受到他人的欺骗，现在我的对手在最近的一次博弈中既不骗人又不受人欺骗，那么我们就真诚地合作；否则我就不客气了。这个策略可以迫使局中人老老实实地再取合作态度。设想一个局中人，如果在他参与的上一次博弈中被对方背叛，那么他在这一次必定应该取背叛策略，如果他不这样去做，他现在的对手根据给定的策略仍会采取“背叛”，因此该局中人不管在这次表示出多好的合作诚意，到了再下一次他参与的博弈中仍然会被人背叛，因此他今天不背叛未必给自己带来好处(注意这就是均衡的含义，在给定其他局中人采取策略剖面中的策略时，局中人若偏离不会为自己带来任何好处)同时，按照规定的策略剖面，如果一个局中人偏离合作，那么不管他今后如何行动，他将永远受到惩罚。基于以上分析，这些策略具有缺乏吸引力的特性，倘若有一个局中人偏离合作、一粒老鼠屎将搅坏一锅粥，引起“全社会”最终拆散为统统背叛的均衡。

§8.5 Pareto 完美均衡与重新谈判检验

我们在本章前面部分已数次提及重复博弈中的重新谈判问题，既然有重新谈判，那么谈判又是怎么回事呢？

回想一下，我们所讨论的均衡其实常常体现为局中人之间的谈判，以及在每一个周期的开头局中人有机会再谈判的结果，于是，通过“偏离将触发一个惩罚均衡”来进行威胁以强制实施“好”结局的均衡可能使人不信任，因为一个局中人可能偏离，然后他提出放弃惩罚均衡而代以另一个均衡，哪个均衡可以使大家的日子更好过一些，我们就称此类均衡限制为 Pareto 完美(Pareto perfection)，Pareto 完美实际上将均衡的 Pareto 最优性与子博弈完美的逻辑结合在一起。原先一种朴素的思想，即在任何子博弈中，局中人总不喜欢采取这样的均衡：从 Pareto 意义上，它是一个劣策略均衡。Pareto 完美推广了上述思想，在子博弈中尽可能地追求 Pareto 最优。

在子博弈中的 Pareto 最优性的约束可以解释为局中人原来协议“重新谈判”的结果，因此也称作“重新谈判验证”(renegotiation-profness)。

对 Pareto 完美与重新谈判验证这两个概念的初步介绍，使我们认识到它们在重复博弈中的重要意义，尤其我们关心，即使局中人偏离了原先谈判所规定的行动，将来的重新谈判是否会有有效的结局。

1. 有限重复博弈中的 Pareto 完美

若从一个例子开始考虑(见图 8.11)。

图 8.11 所示的盈利矩阵实际上是在 §8.1 中研究有限重复博弈时，除了人为地添加(Q_1,Q_2)这个 Nash 均衡外，又人为地添加了两个 Nash 均衡结局(P_1,P_2)与(W_1,W_2)。现在我们有了四个

纯策略 Nash 均衡：(L_1,L_2)，(Q_1,Q_2)，(P_1,P_2)与(W_1,W_2)，分别具有盈利向量(1,1)，(3,3)，(4,1/2)与(1/2,4)。以(L_1,L_2)与(Q_1,Q_2)相比，毫无疑问地，局中人喜欢(Q_1,Q_2)而不喜欢(L_1,L_2)，因为(Q_1,Q_2)Pareto 优于(L_1,L_2)。更为重要地，在图 8.11 中，不存在 Nash 均衡(x,y)，使得局中人一致地喜欢(x,y)，而不喜欢(Q_1,Q_2)、(P_1,P_2)和(W_1,W_2)。我们称(Q_1,Q_2)、(P_1,P_2)与(W_1,W_2)位于图 8.11 阶段博弈的 Nash 均衡盈利集的 Pareto 领域(Pareto frontier)。

局中人 2

		L_2	R_2	Q_2	P_2	W_2
	L_1	1,1	5,0	0,0	0,0	0,0
	R_1	0,5	4,4	0,0	0,0	0,0
局中人 1	Q_1	0,0	0,0	3,3	0,0	0,0
	P_1	0,0	0,0	0,0	4,1/2	0,0
	W_1	0,0	0,0	0,0	0,0	1/2,4

图 8.11

假定图 8.11 阶段博弈重复两次，在第 2 次博弈开始之前观察到第一周期的结局，进而假设局中人预测第 2 次周期的结局如下(同样地，可以看作一种谈判协议)：

如果第 1 次结局是(R_1,R_2)，则为(Q_1,Q_2)；

如果第 1 周期结局是(R_1,y)，y 是除 R_2 之外的任何策略(局中人 2 发生偏离)，则为(P_1,P_2)；

如果第 1 周期结局是(x,R_2)，x 是除 R_1 之外的任何策略(局中人 1 发生偏离)，则为(W_1,W_2)；

如果第 1 周期结局是(x,y)，其中 x 是除 R_1 之外的任何策略，y 是除 R_2 之外的任何策略(两人同时偏离)，则为(Q_1,Q_2)。

我们利用上述预期得到一次性博弈盈利矩阵如图 8.12。

	L_2	R_2	Q_2	P_2	W_2
L_1	4,4	5.5,4	3,3	3,3	3,3
R_1	4,4.5	7,7	4,0.5	4,0.5	4,0.5
Q_1	3,3	0.5,4	6,6	3,3	3,3
P_1	3,3	0.5,4	3,3	7,3.5	3,3
W_1	3,3	0.5,4	3,3	3,3	3.5,7

图 8.12

图 8.12 显示：$((R_1,R_2),(Q_1,Q_2))$是重复博弈的子博弈完美，$((Q_1,Q_2),(Q_1,Q_2))((L_1,L_2),(L_1,L_2))$也是子博弈完美，但显然$((R_1,R_2),(Q_1,Q_2))$Pareto 优于$((L_1,L_2),(L_1,L_2))$，$((Q_1,Q_2),(Q_1,Q_2))$。这里的$((R_1,R_2),(Q_1,Q_2))$是 Pareto 完美的。

将图 8.12 与图 8.5 相比，在图 8.5 中的两阶段重复博弈，局中人在第一周期发生偏离后，惩罚的唯一方法是(L_1,L_2)（(Q_1,Q_2)是奖励不是惩罚），这是一个在 Pareto 意义下的劣均衡，可怜的惩罚者在惩罚他人的时候自己也同时遭殃。但在图 8.12 这个例子中不会出现此类尴尬与困难，这里有三个均衡位于 Pareto 领域，一个是(Q_1,Q_2)，它可以用于奖励两个局中人在第一周期的良好表现（当然，在大家犯“错误”时，也可以协商在第二周期取(Q_1,Q_2)作为补偿）。另外二个，(P_1,P_2)与(W_1,W_2)，它们用来惩罚单方面偏离者，而且在惩罚的同时还奖励了惩罚者，于是，如果提倡在第二周期进行惩罚，那么不存在惩罚者会喜欢阶段博弈其他的均衡，这里，惩罚者不能被说服去重新谈判惩罚问题。

那么，Pareto 完美究竟是怎么回事呢？我们仅从纯策略均衡角度考虑。

设上例中阶段博弈为 g，G^T 表示 T 次重复博弈，（在上例中 $T=2$）令 P^T 表示 G^T 的纯策略子博弈完美均衡的盈利集合（在上例中，依假设，$P^T=\{(7,7),(4,4),(6,6)\}$），令 $Q^1=P^1$ 和 $R^1=$

$Eff(P^1)$。

这里要对 $Eff(\cdot)$作出说明，对于 n 维欧氏空间 R^n 中的任意集合 C，$Eff(C)$表示 C 中强有效点的集合，即，这些点 x 满足如下条件：$x \in C$，那么在 C 中找不到一点 $y(y \neq x)$使得 $y \geqslant x$。举例来说，在上例中，$R^1 = Eff(P^1) = \{(3,3),(4,1/2),(1/2,4)\}$，这是因为 $P^1 = \{(1,1),(3,3),(4,1/2),(1/2,4)\}$，$R^1$ 应从 P^1 中去掉(1，1)而得到，因为$(1,1) < (3,3)$。

对于 $T > 1$，令 $Q^T \subseteq P^T$，表示这样的纯策略于博弈完美均衡盈利的集合，这些盈利在博弈的第二周期中可以用 R^{T-1}中的持续盈利来实施，且令 $R^T = Eff(Q^T)$。

G^T 的一个完美均衡 σ，称为 Pareto 完美，必须满足如下条件：对每一个时间 t 和历史 h^t，在 σ 下的持续盈利在 R^{T-t}中。

根据定义，Pareto 完美首先是一个完美均衡，而且在任何子博弈的部分，其相应的盈利必是强有效的。因此，我们说它将 Pareto 最优与子博弈完美优点结合在一起。

在上例中，应用 Pareto 完美的上述定义，G^2 的子博弈完美均衡$\{(R_1,R_2),(Q_1,Q_2)\}$，对 $t=1$(因为 $T=2$，t 只能取 0 与 1)与历史 $h^1 = (R_1,R_2)$，持续盈利$(3,3) \in R^1$，而 $R^2 = (7,7)$为唯一，故它为 Pareto 完美。

现再举一个较典型的例子：

例 8.4 考虑图 8.13 所示的阶段博弈。

	b_1	b_2	b_3	b_4
a_1	0,0	2.4	0.0	5.5,0
a_2	4.2	0,0	0,0	0.0
a_3	0,0	0,0	3,3	0.0
a_4	0,5.5	0,0	0,0	5,5

图 8.13

图 8.13 中的阶段博弈有三个纯策略 Nash 均衡(a_2,b_1)，(a_1,b_2)，(a_3,b_3)，更有效的 (a_4,b_4) 不是 Nash 均衡。因此 $R^1=\{(4,2),(3,3),(2,4)\}$。考虑两次重复博弈 G^2，设 $\delta=1$，局中人充分地有耐心。与前面那个例子相似，因为 R^1 有 3 个元素，我们可以在第一周期实现盈利(5,5)：

第一周期的结局是(a_4,b_4)，则第二结局为 (a_3,b_3)。

如果第一周期的结局是 (a_4,y)，y 为非 b_4 的任意策略，则第二周期为 (a_2,b_1)。

如果第一周期的结局是 (x,b_4)，x 为非 a_4 的任意策略，则第二周期为 (a_1,b_2)。

如果第一周期的结局是(x,y)，$x\neq a_4$，$y\neq b_4$，则第二周期为(a_3,b_3)。

此时 $P^2=\{(7,5),(5,7),(6,6),(8,8)\}$，显然 $R^2=(8,8)$。因此 G^2 有唯一 Pareto 完美盈利。$\{(a_4,b_4),(a_3,b_3)\}$是 G^2 的唯一 Pareto 完美均衡。现在考虑阶段博弈重复三次 G^3，由于 R^2 只有单点(8,8)，根据 Pareto 完美的定义，我们无法在第一个周期强行实施有效结局(a_4,b_4)，因为根据子博弈完美的要求，在持续博弈 G^2 中不可能有赏罚分明的余地。于是 Pareto 完美要求在第一周期发生静态博弈的 Nash 均衡。$Q^3=R^3=\{(12,10),(11,11),(10,12)\}$，即将 3 个 Nash 均衡的盈利加到(8,8)上去。在 Q^3 中有 3 个元素可变化，也就是说，在持续博弈 G^3 中存在着奖励或惩罚的可能，使得如果重复阶段博弈 4 次的话，在第一周期又可以实施有效结周(a_4,b_4)了。这些想法可以一直延续下去，Benoit 与 Krishna 证明了集合 R^T 交替地，当 T 为奇数时有 3 个元素，当 T 为偶数时只有 1 个元素。当 $T\to\infty$时，R^T/T(即每周期的平均盈利)趋于(4,4)。然而，Benoit 与 Krishna 于 1985 年证明了当 T 相当大时，有效盈利(5,5)可以在一个子博弈完美均衡中近似地得到。也就是说，当 T 适当大时，存在一个子博弈完美均衡，其盈利肯定大于同样

的 T 时 G^T 的 Pareto 完美均衡的盈利。这个事实表明了，Pareto 完美均衡虽然一定是子博弈完美均衡(Pareto 完美的定义所要求)，但是它未必在所有的子博弈完美均衡中都是 Pareto 有效的，这是因为对于 Pareto 完美持续博弈的限制减少了引导局中人采用有效策略(a_4,b_4)的频率。

关于有限重复博弈的 Pareto 完美均衡的定义，有各种各样的提法，我们仅举一些略加介绍。读者若希望了解更多有关知识，可以查阅有关文献。

2. 无限重复博弈中的重新谈判检验(renegotiation-proofness)

我们在阐述有限重复博弈中的 Pareto 完美概念时，常从最后一次博弈出发，使子博弈的结局盈利为 Pareto 最优，然后往前确定前面的行动，这有点后退归纳的味道。这种定义的方法当然不适合于无限重复博弈。关于无限重复博弈中的重新谈判检验或 Pareto 完美的定义方式看来必须另起炉灶，目前有若干不同的定义互相"竞争"，我们先介绍最易处理的方式之一，它是由 Farrell 与 Maskin 于 1989 年提出，称为无限重复博弈中的"弱重新谈判检验"(简记为 WRP)，WRP 的出发点是，存在一个由"外因"选取的可能均衡盈利集合 Q，Q 在任意 t 及历史 h^t 是可想象得到的，Q 中的每一个盈利必须要求只能对应于 Q 中其他均衡的持续盈利。具体地说，如果令 $c(\sigma,h^t)$为给定历史 h^t 下由 σ 产生的持续盈利，再令 $c(\sigma)=U_{t,h^t}c(\sigma,h^t)$为策略剖面 σ 的所有持续盈利的集合。如果 $v\in Q$，那么一定存在一个子博弈完美均衡 σ，它具有盈利 v 而且 $C(\sigma)\subseteq Q$。据此，我们可以根据一些子博弈完美均衡来"人为"地构造这样的 Q。如果如此构造的 Q 中，没有一个均衡盈利 Pareto 劣于 Q 中的另一个均衡，那么我们称集合是 WRP。

我们举例来进一步阐述 WRP 思想。假如阶段博弈为囚徒窘境，其盈利矩阵如图 8.14。

		局中人 2	
		C	D
局中人 1	C	2,2	−1,3
	D	3,−1	0,0

图 8.14

C 表示合作、D 表示背叛，(D,D)是唯一的静态均衡，众所周知，“总是背叛”是无限重复博弈的子博弈完美均衡，以(0,0)构造单点集 Q，因为“总是背叛”的一切持续盈利都是(0,0)，可见，一直(D,D)下去，是 WRP。不难设想，阶段博弈的任何静态均衡都具有上述特性。

但是，同样也是子博弈完美的触发策略：最初合作，一旦有人偏离则永远回到静态均衡，却不是 WRP。理由很简单，对应于“合作状态”策略的盈利当然 Pareto 优于惩罚状态的盈利。就是说，由于一开始采取合作策略，因此“总是合作”的盈利包含在可能“协议”的集合 Q 中，于是局中人总是重新谈判以便从无休止的惩罚回到合作状态，这样对双方都有利一些。

如果策略确定为“在第一周期取 C，在以后的周期中，如果上一周期的结局是(C,C)或(D,D)，那么取 C；如果上一周期的结局是(C,D)或(D,C)，那么取 D”，我们称该策略为“完美针锋相对”，读者可注意到，它与针锋相对策略有所不同。这个策略剖面不是 WRP，因为在发生单方面偏离之后紧接的下一个周期中，不理睬偏离而采取(C,C)将会更有效。关于这一点只要计算一下从该周期开始的两个持续盈利就可以了。

那么“总是合作”是不是 WRP 呢？因为我们至少知道，只要局中人有充分的耐性，“总是合作”是子博弈完美均衡。1989 年 Farrell 与 Maskin，以及 Damme 证明了，如果折扣因子充分地接近于 1，“总是合作”是 WRP 结局。其实，他们证明了在无限重复囚徒窘

境中，无名氏定理以重新谈判检验的形式，其结论仍成立。持别地，下述"赎罪"策略是 WRP 并且具有有效盈利。策略定义为"从双方都取 C 开始。如果单个局中人 i 偏离至 D。则转向对 i 的惩罚状态，在该状态中，局中人 i 取 C 而另一个局中人取 D，直到局中人 i 第一次取 C 而此刻博弈回到合作状态才结束惩罚状态。"为验证这个策略剖面是 WRP，首先考察它是否子博弈完美均衡。要做到这一点，只要验证如下几个方面就可以了。在合作状态，任何一个局中人偏离将不会为它带来好处；一旦局中人 1 偏离，它必定会为了极大化自己盈利而在受到惩罚之后回到合作状态；如果局中人 1 偏离，局中人 2 采取惩罚才对自己有利。现在我们分别研究这三种情况：

(1)如果(假设)局中人 1 偏离合作状态，那么它将承受一个周期的惩罚，倘若折扣因子充分地接近于 1 的话，对于局中人 1 来说，偏离合作不是一个合意的策略。因为若不偏离，两个周期的盈利和应当为 $2+2\delta$，若偏离，他先获得 3，然后受到一个周期的惩罚，这两个周期的盈利和为 $3-\delta$，当 δ 比较大时，显然有 $2+2\delta>3-\delta$。

(2)一旦局中人发生偏离，如果它甘受惩罚并在之后回到合作状态，那么它的平均持续盈利应为

$$(1-\delta)\left(-1+\frac{2\delta}{1-\delta}\right)=-(1-\delta)+2\delta=3\delta-1>0 \quad \left(\delta>\frac{1}{3}\right) \tag{8.71}$$

如果局中人 1 偏离策略剖面的给定，譬如它坚持再取 D 一次，那么它面临的结局当为(D,D)，(C,D)与永远的(C,C)，其相应的平均持续盈利为

$$\begin{aligned}&0-\delta(1-\delta)+(1-\delta)\cdot\frac{2\delta^2}{1-\delta}\\&=\delta[-(1-\delta)+2\delta]\\&=\delta(3\delta-1)<3\delta-1\end{aligned} \tag{8.72}$$

只要 $\delta \neq 1$，坚持背叛还不如老老实实地回到合作策略上来。

(3)当局中人 1 因偏离而受惩罚时，局中人 2 的盈利为

$$(1-\delta)(3+\frac{2\delta}{1-\delta})=3(1-\delta)+2\delta=3-\delta \qquad (8.73)$$

若局中人 2 不去惩罚局中人 1，而局中人 1 则在背叛之后立即回到合作，此时局中人 2 的平均持续盈利为

$$(1-\delta)\cdot\frac{2}{1-\delta}=2 \qquad (8.74)$$

显然 $3-\delta>2$（只要 $\delta\neq 1$ 就行）可见，为极大化自身盈利，局中人 2 必定会按照策略剖面的规定去惩罚局中人 1。

综上所述，“赎罪”策略是子博弈完美均衡。

现在我们试图证明它是 WRP，这是件非常简单的事情。注意到对任意时刻 t，历史 h^t 无非有三种情况：要么是合作状态；要么局中人 1 偏离；要么局中人 2 偏离。对应于这三种子博弈的（平均）持续盈利分别为(2,2)，$(3\delta-1,3-\delta)$与$(3-\delta,3\delta-1)$，它们之中没有一个 Pareto 劣于其他一个，因此按 WRP 的定义，这个策略剖面是 WRP。

在无限重复囚徒窘境中得到上述有效 WRP 盈利的关键是使用策略组合(C,D)惩罚了局中人 1，而惩罚者从中又得到了奖励。在其他的某些博弈中，在奖励局中人 2 与惩罚局中人 1 之间可能存在着交易，这可以阻止整个有效个体理性盈利集成为 WRP。我们在这里不准备讨论。

既然有弱重新谈判检验，想必也会有强重新谈判检验(SRT)概念出现。很自然地，Q 称为强重新谈判检验。它首先必须是 WRP，由于 WRP 集合里常有不止一个 WRP，如果不存在这样的 WRP 集，它具有的盈利严格地 Pareto 优于 Q 中任意盈利，那么将 Q 称为 SRT 也许是比较贴切的。不幸的是，这样的强重新谈判检验未必存在。因此，本书暂不展开这方面的讨论。

§8.6 公共信息不完美时的重复博弈

我们在前面所研究的重复博弈模型有一个明显的特点，那就是可观察行动的，每一个周期的阶段博弈结束时，每个局中人都可以观察到其他局中人的现实行动。在经济学研究中，情况并不总是那样地美好与理想，经常地，局中人能够接受到的信息只不过是对手在阶段博弈中策略的不完美信号。经济学家主要关心公共信息博弈，在每个周期的最后，所有局中人观察到一个"公共结局"，它与阶段博弈的行动向量相关联，并且每个局中人的已获盈利仅仅依赖于它自身的行动以及公共结局。于是，对手的行动仅通过对公共结局分布的影响而影响到局中人的盈利。

Green 与 Porter 于 1984 年对此类博弈模型进行研究，试图解释"价格战"的发生，该工作部分地受到 Stigler(1964)研究的启发。在 Stigler 模型中，每家公司只观察到自己的销售情况而无法知道其对手的价格或产量。而消费者需求的总水平是随机的(现在，大概不会有人反对这个假设)，于是，公司销售额的下降可能要么归因于需求水平的下降，要么是由于未能观察到的某个对手的削价而引起。因为每家公司关于其对手行动的唯一信息就是它自己的已销售水平，没有一家公司知道对手看到了什么，关于已采取的行动也没有公共信息。Green-Porter 模型则假设了博弈具有公共信息，这使得模型更便于分析。在他们的模型中，每家公司的盈利依赖于自已的产量和共同观察到的市场价格。公司观察不到他人的产量，市场价格依赖于未观察到的需求冲击，也依赖于总产量。因此，一个意想不到的低市场价格可能归因于某对手的未意料的高产量或者意料不到的低需求。

不完美公共信息的重复博弈例子还有许多，例如"喧闹的"囚徒窘境，其中，局中人有时候不经意地选择了"错误"行动，使得观

察到的行动仅是打算实行的一个不完美的信号。

上述博弈以及类似的有关博弈都是“重复道德风险”的例子。鉴于篇幅，我们不在这里作专题描述。

1. 触发价格策略(trigger-price strategies)

在 Green 与 Porter 关于垄断模型的分析中，集中于触发价格策略.首先建立一个简单的模型。

阶段博弈为 n 个局中人同时从自己的行动空间中选择相应的策略，公司 i 选择自己的产量 $a_i\in[0,\overline{Q}]$。共同观察得到的结局是市场价格 y，y 所属的价格空间 Y 显然可假设为实数空间的子空间。每一个行动剖面 $a\in \mathop{X}\limits_i[0,\overline{Q}]$导致 Y 上的一个概率分布——公共信息具有不确定性，不妨以 $\pi_y(a)$表示在 a 下取 y 的概率，$\pi(a)$表示由 a 导致的 Y 上的概率分布。Green 与 Porter 假设 $\pi(a)$仅依赖于公司产量之和 $\sum_{i=1}^n a_i$，并且在每一个行动剖面下，每一个价格有正概率。局中人 i 的已获盈利 $r_i(a_i,y)$假定独立于其他局中人的行动，因此在策略剖面 a 下，局中人 i 的期望盈利为

$$g_i(a)=\sum_y \pi_y(a)r_i(a,y) \tag{8.75}$$

在重复博弈中，t 周期开始时的公共信息为

$$h^t=(y^0,y^1,\cdots,y^{t-1}) \tag{8.76}$$

本模型中，假设盈利函数为对称的且限于局中人在每周期选择相同行动的均衡，即对所有 t 及 h^t，当 $i\neq j$ 时应有

$$\sigma_i(h^t)=\sigma_j(h^t)$$

也就是说，均衡是强对称的。

触发价格策略剖面有两个可能状态：一为合作状态，所有公司生产相同的产量，不妨记作 $\hat{a}$。只要每一个周期的实际价格 y^t 至少是“触发价格”$\hat{y}$。那么公司就继续留存于合作状态。如果 $y^t<\hat{y}$，那么就转向 $\hat{T}$ 个周期的“惩罚状态”。在惩罚状态中，每一个周期局中人取静态 Nash 均衡策略 a^*，不管实际结局 y 如何；结束了 $\hat{T}$

个周期的惩罚之后，重新回到合作状态。

触发策略其实是指各公司在产量上合作以垄断市场价格，至少不会小于一个最低的垄断价格 $\hat{y}$。如果发生某局中人偏离合作擅自提高产量或者由于市场需求量减少而引起市场价格低于 $\hat{y}$，那么在随后的各局期内各公司进行理性竞争，$\hat{T}$ 个周期后重新在产量上进入合作状态。

显然，在触发价格策略中，存在着三个未知参数：$\hat{a}$，$\hat{y}$ 与 $\hat{T}$。它们将直接影响到触发价格策略是否为均衡。

首先我们会关心触发价格均衡是否存在。简单地作一个设想，一个特殊的触发价格策略为，在每一个周期中干脆令 $\hat{a}=a^*$，即各公司均取竞争博弈的静态均衡策略(例如 Cournot 竞争博弈，静态均衡策略的产量对各公司是相等的)，众所周知，在重复博弈中，每个周期取静态均衡策略构成子博弈完美均衡，因此，这足以说明触发价格均衡的存在性没有任何问题。那么一般情况的触发价格均衡应该如何刻画呢？我们在下面进行详细讨论。固定 $\hat{a}$ 与 $\hat{y}$，即各公司互相之间确定一个最低市场价格以及规定各自产量都等于 $\hat{a}$，令

$$\lambda(\hat{a})=\text{Prob}(y'\geqslant\hat{y}\,|\,\hat{a}) \tag{8.77}$$

条件概率 $\lambda(\hat{a})$ 是指，局中人都生产相同量 $\hat{a}$ 产品时，共同观察到的市场价格大于等于触发价格的概率。如果所有的局中人都遵照策略的话，那么标准化盈利为(这里，标准化是指大家采取静态均衡产量 a^* 时盈利“标准化”为 0。这样，使我们在计算方面非常便利，从下面的(8.78)式中立即可以看到标准化的优点)：

$$\hat{v}=(1-\delta)g(\hat{a})+\delta\lambda(\hat{a})\hat{v}+\delta(1-\lambda(\hat{a}))\delta^{\hat{T}}\hat{v} \tag{8.78}$$

由(8.78)式，不难得到

$$\hat{v}=\frac{(1-\delta)g(\hat{a})}{1-\delta\lambda(\hat{a})-\delta^{\hat{T}+1}(1-\lambda(\hat{a}))} \tag{8.79}$$

(8.78)式右端后两项之和为 $\delta\hat{v}$ 与 $\delta^{\hat{T}+1}\hat{v}$ 的加权平均。因此小于

$\max(\delta\hat{v},\delta^{\hat{T}+1}\hat{v})\leqslant\delta\hat{v}$，于是 $\hat{v}\leqslant g(\hat{a})$ 为显然事实。要考虑均衡，自然要考虑持续盈利的极大化问题。由(8.79)式，$\hat{v}$ 在两种情况下可达到 $g(\hat{a})$，一是如果 $\lambda(\hat{a})=1$，这使得只要所有局中人保持一致，惩罚的概率等于 0；另一个是 $\hat{T}=0$，惩罚的长度等于 0。惩罚长度等于 0 就是没有惩罚周期。它仅当 $\hat{a}$ 为静态均衡产量时方有可能。在这种情况下，不需要通过惩罚来促进激励，此时已知策略构成子博弈完美。即使 $\hat{a}$ 不取为静态均衡产量 a^*，使 $\hat{v}=g(\hat{a})$ 的还有一个可能是 $\lambda(\hat{a})=1$，即除非有人偏离，否则不存在惩罚的可能。这种情况是否可能呢？如果局中人的行动(即公司的产量)完全被观察到的话。产量全在“监督”之中，没有一家公司主动违规从而招致惩罚。

以上两种情况均使 $\hat{v}=g(\hat{a})$。由于 $\hat{T}=0$ 的情况只有在 $\hat{a}=a^*$ 时成立，因此不必继续讨论。我们关心的是 $\lambda(\hat{a})$ 可能小于 1 的情况。因为在 Green-Porter 模型中，有一个“全支撑”假设，即对所有 a 及所有可能市场价格 y，$\pi_y(a)$ 都取正数，或者干脆说，不管 a 取什么，任何市场价格 $y\in Y\subseteq R$ 都是存在着可能性的。全支撑假设至少告诉我们，$\lambda(\hat{a})$ 存在小于 1 的可能。于是，在该模型中，“只要没有局中人偏离合作就不发生惩罚”的唯一触发价格策略也具有如下性质：在任意系列的市场价格结局之后决不发生惩罚（读者应注意这里的描述与 $\lambda(\hat{a})=1$ 的描述之间的差异。这种描述之间的差异其原因正在于“全支撑”假设）。因此，决不发生惩罚的唯一的触发价格均衡一定是那些存在重复实施静态均衡现象的行动系列。于是，如果均衡是对静态盈利作出改善，那么它的 $\lambda(\hat{a})$ 将一定比 1 小，或者说，一定存在惩罚，而对偏离合作行为给予严厉的惩罚也必然存在着成本或损失。特别地，对于固定的 $\hat{y}$ 与 $\hat{a}$，随着惩罚长度 $\hat{T}$ 的增加，惩罚成本随之增加，均衡盈利将减少。

上述分析似乎是说，为了减少惩罚成本，最优的办法是适当缩减惩罚长度 $\hat{T}$，但是，事实上，非常长的甚至无限的惩罚也许是最

优的，因为通过增加惩罚长度，有可能减少触发价格，或者增加合作状态时的盈利。最优触发价格均衡将使(8.79)式中 $\hat{v}$ 在约束条件下极大化，其激励约束为，没有一个局中人通过偏离合作状态而获利，用数学公式来表示，设局中人 i 偏离 $\hat{a}$，对一切 a_i，相应盈利应为

$$\hat{v}^*=\frac{(1-\delta)g(a_i,\hat{a}_{-i})}{1-\delta\lambda(a_i,\hat{a}_i)-\delta^{\hat{T}+1}(1-\lambda(a_i,\hat{a}_i))} \tag{8.80}$$

$\hat{v}^*\leqslant\hat{v}$ 意味着对一切 a_i 成立如下不等式

$$\begin{aligned}&(1-\delta)g(\hat{a})[1-\delta\lambda(a_i,\hat{a}_{-i})-\delta^{\hat{T}+1}(1-\lambda(a_i,\hat{a}_i))]\\&\geqslant(1-\delta)g(a_i,\hat{a}_{-i})[1-\delta\lambda(\hat{a})-\delta^{\hat{T}+1}(1-\lambda(\hat{a}))]\end{aligned} \tag{8.81}$$

(8.81)式也可整理为

$$\begin{aligned}&(1-\delta)[g(a_i,\hat{a}_{-i})-g(\hat{a})]\\&\leqslant\frac{\delta(1-\delta^{\hat{T}})[\lambda(\hat{a})-\lambda(a_i,\hat{a}_{-i})](1-\delta)g(\hat{a})}{1-\delta\lambda(\hat{a})-\delta^{\hat{T}+1}(1-\lambda(\hat{a}))}\end{aligned} \tag{8.82}$$

对一切 a_i 成立。

因此，从公司的观点出发，最优触发价格均衡中的 $\hat{a}$、$\hat{T}$ 与 $\hat{y}$，应当满足在约束条件(8.81)式下使(8.79)式极大化。

在触发价格均衡中，博弈最终进入惩罚状态这一事件具有概率 1，不严格地说，这一点与“价格战”(price-wars)的想法一致，可是应看到在触发价格均衡中所有局中人准确地预测他们的对手将决不偏离。于是，通过推断某些公司在前面周期选取了高产量不会触发价格战。相反地，所有局中人正确地假设他们的对手在上一周期选择了合作产量，市场价格低是因为需求冲击的缘故，而不管怎样，“惩罚”是作为对低水平的现行要求的自身反应。

第三部分　不确定结局的博弈问题

第九章　道德风险问题

博弈问题分为确定性与不确定性两种，我们在前面介绍与研究的博弈大多数属于确定性的，即所有局中人肯定知道任何策略剖面的结局。然而大千世界经常出现复杂得多的情况，例如，公司很难准确地知道当降低产品的销售价格时，将会多出售多少货物。股民不知道他们交易的股票上市公司的变现值，从而使他们在变幻莫测的股市博弈中承受一定风险。在市场竞争博弈中，局中人无法确定策略剖面的结局，因此，这类博弈问题的研究更具复杂性。本章以保险市场与就业市场为例对不确定性的博弈稍作介绍。

§9.1　道德风险与不完全保险

构造一个简单的失业保险模型：

某甲为一个(假设的)公司服务，由于业务原因该公司的人员流动性较大。在每月的1日，公司决定该月雇佣的人数。有时候，公司需要裁员并解雇若干在上个月还在为它工作的人员，有时候公司将增加雇员，召回上月被解雇的人，且登广告招收新人员。因

为某甲的学历较低，职位不高，在每个月里他面临被解雇的可能为50%。现在作一个规则性假设，解雇仅实施一个月，公司的政策是在下一月份里可能召回任何被解雇的工作人员，或者说，甲如果在这个月被解雇，那么他在下个月里仍有50%的可能回公司上班。假如某甲遭到解雇，他不可能立即找到工作。在我们的模型中，假设他不可能有机会找到一个月的其他工作。又设如果甲被雇用，则每月报酬为1 000元，就是说，每月甲有50%机会挣1 000元，也有50%机会失业在家而两手空空。解雇概率为50%这个事实是众所周知的，且甲无能力对它进行控制。

假定甲的月收入为 Y，为方便起见，单位以千元计。甲的效用函数可记为 $U(Y)$，U 是月收入的递增函数，但递增的速率是随 Y 的增加而减少，倘若 U 是 Y 的充分光滑函数的话，上述说法意味着其一阶导数为正、二阶导数为负。如果某甲不参加任何失业保险，那么每个月他的期望效用为

$$U_0=0.5U(0)+0.5U(1) \tag{9.1}$$

这实际就是所谓保留效用。

每月月底，甲可以从保险公司(简记为乙)处购买失业保险单(unemployment insurance policy)，如果下一个月甲被解雇，保险单将补偿他的部分月收入，以 I 记补偿的金额。当然保险公司不会单方面做亏本买卖。作为一种交换，乙要求甲支付不能退还的保险费(preminum)，其金额大小记作 p。因此，p 与 I 完全确定了失业保险单的内容。不妨以向量 (P,I) 表示之。

现在假设甲购买了保险单 (P,I)，很容易计算他在购买后的下一个月的期望效用：

$$EU(P,I)=0.5U(I-P)+0.5U(1-P) \tag{9.2}$$

显然，只有在 $EU(P,I)\geqslant U_0$ 时，甲才会去购买这份失业保险。这仅是博弈的一方所考虑的效用，必须考虑博弈的另一方——乙(保险公司)的期望效用，为方便起见，本模型不考虑乙将支付的

管理成本。乙从甲处收取保险费 P，而支付赔偿金 I 的概率等于甲下月失业的概率，因此乙的期望效用函数为

$$ER(P,I)=P-0.5\times I \tag{9.3}$$

如果 $ER(P,I)=0$，从保险统计角度，平均而言，保险公司既不赚顾客（这使顾客愿意买保险）又不亏本（这是保险公司生存的基本点），无疑这是一个公正的准则。因此，当且仅当 $P=0.5\times I$ 时，保险单是公正的。

注意到在本博弈模型中，我们讨论的是期望效用，在以前引进混合策略概念时，也讨论过期望盈利或效用。读者不难发现在混合策略中，概率是受到局中人控制的，局中人为使自己获得极大化盈利而主动设计采取行动的概率。在我们的失业保险博弈中，不确定的结局并不为局中人所能控制，局中人为极大化自己的获益只能利用关于概率的知识或统计估计从而设计公正合理的保险单 (P,I)。

在公正保险单 $(0.5I,I)$ 下，甲的期望效用为

$$EU(0.5I,I)=0.5U(0.5I)+0.5U(1-0.5I) \tag{9.4}$$

欲使（9.4）式达到最大，一阶条件为

$$U'(0.5I)=U'(1-0.5I) \tag{9.5}$$

由于 $U''<0$（U 为凹函数），（9.5）式蕴含着 $0.5I=1-0.5I$，或者 $I=1$。结论很清楚，如果甲的效用函数为递增且凹，那么最优的选择是购买全保险的保险单（0.5，1），即他支付 500 元保险费，一旦他下个月被解雇，他将获得1 000元赔偿。也就是说，无论他在下个月是否被解雇，他下月的收入是 500 元。这个结论很有趣，购买全保险的决策并不依赖于他的效用函数的形式。

在上述简单失业保险模型里，买卖保险的双方在交易前后，解雇概率或解雇之后的损失等各种因素不发生变化，在保险统计角度公正的情况下，双方以公正的保险单进行交易并按照此规则行事，不存在所谓“道德风险”问题。但是，如果下述两个条件成立的

话，在失业保险市场就出现道德风险：

(1)在购买保险单后，买者甲采取改变损失大小或改变损失发生概率的行动。

(2)观察这些行动的花费是如此地昂贵，以至监控买者甲的行动在经济上是不可行的。因为监控的高成本，使得保险公司无法根据购买者以后的行动制定所付赔偿的有关条款。

在失业保险市场，这两个条件同时成立。假设甲知道自己已经从目前的岗位上被解雇，他不会消极地等待再下一次被公司召回(这是原来模型的规则)，取而代之，他将主动地寻找另外的工作。他越是努力地去找新工作，那么他找到新工作的概率也就越大，尽管公司可以几乎不花费成本地证实甲的工作状况，可是公司根本不能够控制甲去寻找新工作的努力。当然，甲很清楚自己在这方面的努力，但保险公司不可能指望甲会真实地报告自己寻找新工作的努力程度。其结果是，如果甲遭到解雇，无论他寻求新工作的努力有多大，保险公司乙将只得按照保险单规定支付赔偿。

假定我们可以用某种办法量化寻找新工作的努力程度 H，例如，H 为寻求新工作所必需支付的费用。在甲遭解雇后找到新工作的概率记作 $\pi(H)$。显然 π 是 H 的递增函数，但我们可以合理地假定这种递增的速度是渐降的，当然，倘若甲根本不去找新工作，那么他有一个新工作的概率为 0，即 $\pi(0)=0$，假定

$$\pi(H)=\min\{\sqrt{H},1\} \tag{9.6}$$

(9.6)式表明，$H\leqslant 1$，即我们假定寻找新工作的花费不超过甲的月收入。众所周知，寻找到一件新工作是较复杂的事件。例如，在什么时候能找到，新工作的报酬如何等等。本章只是讨论结局为不确定情况的展开型博弈，因此我们尽可能地使模型简明一些，为方便起见，我们假设如果甲不立即出去积极地应聘并在该月月初的第一天找到新工作的话，那么他在整个月内均处于失业状态。于是，我们仅面临甲的两种可能的工作状态：要么他在整个月都工作

并具有月收入 1(千元)，要么他在整个月内均处于失业状态而无收入进账。

现在，甲的效用函数将明显地不同于前面所介绍的不出现道德风险时的情况，由于寻找新工作需要花费，U 不仅依赖于月收入 Y，同时也依赖于 H。为方便以及今后计算的具体化起见，我们取甲的效用函数为下面的形式：

$$U(Y,H)=\log_2(Y)-H \tag{9.7}$$

这里记号 $\log_2(Y)$表示以 2 为底的对数。

全失业保险，在甲被解雇的那个月内提供给他的赔偿相当于他出去找到另外工作的收入，这使得他找到新工作所获纯利立即减少到 0。假如甲买了全失业保险，那么当他被解雇时他将不会去找新工作，因而在那个月毫无疑问不会去工作，其结果，甲在那个月内真正(或事实上的)失业的概率为 50%，另一方面，如果甲所购买的保险单仅提供他相当于部分月收入的赔偿，使得他感到有必要去寻新工作。这样，甲在任何一个确定的月份里处于失业状态的概率将可能减少为

$$\begin{aligned} P\{\text{某月甲失业}\} &= P\{\text{某月甲被公司解雇}\}\cdot \\ &\quad P\{\text{甲找不到新工作}\mid\text{甲被公司解雇}\} \\ &=0.5\cdot(1-\pi(H)) \end{aligned} \tag{9.8}$$

显然，这个概率将影响到保险公司关于任何形式失业保险的期望收益或损失。因此保险公司在对保险单进行定价之前，必须往前看并预知甲在购买保险之后的所作所为，这一点与其他商品买卖是相当不同的。

这里展开的博弈实际上是委托一代理博弈，保险公司乙是委托人(principal)，某甲是代理人(agent)。在这个博弈中有 5 个行动：

(1)保险公司首先行动，公布它的保险费一览表 $P(I)$ —— 对每一种赔偿要求所索取的保险费。

(2)某甲决定购买多少保险。

(3)"自然"行动以确定甲是被解雇还是继续留用，如果甲未被解雇，则博弈结束。要不，甲将在第四步采取行动。

(4)甲再次行动，选择如何努力去寻找一个新工作。

(5)"自然"作最后一次行动，确定甲是否找到新工作。如果甲找到新工作并求任，保险公司则不必做任何事(由于甲仍有1千元收入，乙不必赔偿)。如果甲找不到新工作，保险单将起作用，乙必须按照协议支付给甲一定的补偿。

保险公司的策略由保险费一览表 $P(I)$ 组成。在均衡状态，保险公司必须在每一张卖出的保险单上挣得期望收益为0。

某甲的策略应当由两个分量组成——因为他在这个模型中有两个行动——第一个分量表示每一种可能的保险费清单上的补偿水平，第二个分量表示在甲所选择的每一种可能的保险合同之后，他寻求新工作的努力水平。因此，根据子博弈完美均衡的要求，必须满足：

(1)在给定的已选择好的保险情况下，某甲选择最佳的寻求新工作的努力水平。

(2)某甲关于补偿水平的选择在给定的保险一览表中是最优的，当然我们假定他然后从事于最优的寻求努力。

(3)假定某甲购买了保险单，其中选择的补偿水平与从事寻求新工作的努力水平使甲的期望效用极大化，公司的保险费一览表将允许公司在每一个补偿水平上(平均而言)不亏不盈。

根据以上三点要求，我们试图求出子博弈完美均衡中 P、I 与 H 值。

假定甲买保险单 (P,I)，并且他被解雇，用努力水平 H 寻求新工作，如果甲找到了新工作的话，他的盈利(或效用)是 $\log_2(1-P)-H$；如果甲找不到新工作从而接受保险公司的补偿，此时他的赢利是 $\log_2(I-P)-H$。综合起来，他的期望效用应为：

$$EU(P,I,H)=\sqrt{H}[\log_2(1-P)-H]+(1-\sqrt{H})[\log_2(I-P)-H]$$
$$=\sqrt{H}\log_2(1-P)+(1-\sqrt{H})\log_2(I-P)-H \qquad (9.9)$$

现在先考虑三点要求中的第一点。即甲选择一个适当的 H 以使得(9.9)式达到在(P,I)固定时的极大值,这个 H 值当然是 P 与 I 的函数,不妨记作 $H^*(P,I)$。我们试图求解 H^*:

当 $I=1$ 时,$EU(P,I,H)=\log_2(I-P)-H$,显然 H 越小,期望盈利越大,但是寻求新工作不可能 $H<0$,从而取 $H^*=0$ 时 $EU(P,I,H)$取得极大值 $\log_2(I-P)$。

当 $I\neq1$ 时,在(9.9)式中对 H 求导数获一阶条件:

$$\sqrt{H}=\frac{1}{2}\log_2\frac{1-P}{I-P} \qquad (9.10)$$

如果 $I>1$,则 $(1-P)/(I-P)<1$,$\log_2\frac{1-P}{I-P}$为负数,显然(9.10)式无解,这表明,当 $I>1$ 时不能用一阶条件(9.10)式。回到(9.9)式,期望效用可写作

$$EU(P,I,H)=-\sqrt{H}\{\log_2(I-P)-\log_2(1-P)\}$$
$$+\log_2(I-P)-H \qquad (9.11)$$

由于$\{\cdot\}$内为正,H 与 $\sqrt{H}$ 前均为负号,因此,唯取 $H=0$ 才使 $EU(P,I,H)$达到极大。

当 $I<1$ 时,可考虑一阶条件(9.10)式。但由于先前我们假设 $H\leqslant1$,因此 $\sqrt{H}\leqslant1$。于是一阶条件要求

$$\log_2\frac{1-P}{I-P}\leqslant2 \qquad (9.12)$$

即

$$\frac{1-P}{I-P}\leqslant2^2=4 \qquad (9.13)$$

或

$$I\geqslant\frac{1+3P}{4} \qquad (9.14)$$

因此，当 I 满足$\frac{1+3P}{4}<I<1$ 时，由(9.10)式可得

$$H^*(P,I)=\left[\frac{1}{2}\log_2(\frac{1-P}{I-P})\right]^2 \tag{9.15}$$

那么，当 $I\leqslant\frac{1+3P}{4}$时，H^* 等于什么呢？从(9.10)式的形式，此时 H 可取大于等于 1 的数，由于 H 的限制，至多只能取 $H^*=1$。综上所述，我们得到

$$H^*(P,I)=\begin{cases}0 & 若\ I\geqslant 1\\ \left[\frac{1}{2}\log_2(\frac{1-P}{I-P})\right]^2 & 若\frac{1+3P}{4}<I<1\\ 1 & 若\ I\leqslant\frac{1+3P}{4}\end{cases} \tag{9.16}$$

(9.16)式至少告诉我们，如果甲所购买的保险使他一旦被解雇后从保险公司处获得的补偿大于他的月收入，那么他不会再去寻求新工作。

现在我们来为公司方面考虑。如果公司卖出保险单(P,I)给某甲，那么它不得不付出 I 给甲的唯一时机为甲失去了工作而且没有找到另外的工作，这件事发生的概率公式表示在(9.8)式中，其概率(并结合公式(9.6)式)为

$$\begin{aligned}0.5(1-\pi(H^*))&=0.5(1-\pi(H^*(P,I)))\\&=0.5(1-\sqrt{H^*(P,I)})\end{aligned} \tag{9.17}$$

于是，保险公司乙的期望效用为

$$ER(P,I)=P-0.5(1-\sqrt{H^*(P,I)})\cdot I \tag{9.18}$$

已经指出，在均衡中，保险公司乙必须关于所有合同恰好不亏不盈，即(P,I)必须满足

$$\begin{aligned}P&=0.5(1-\sqrt{H^*(P,I)})\cdot I\\&=0.5(1-\frac{1}{2}\log_2\frac{1-P}{I-P})\cdot I\end{aligned} \tag{9.19}$$

(9.19)式中 I 应满足 $\frac{1+3P}{4}<I<1$。我们不考虑 $H^*=0$ 或 1 的情况，因为 $H^*=0$ 表明甲不去寻找工作，而 $H^*=1$ 表明甲为寻新工作付出的努力恰好等于他的月收入。

如果我们以 p 表示每补偿 1 元钱的保险成本，则 $p=P/I$，即 $P=pI$，(9.19)式成为

$$pI=0.5\left(1-\frac{1}{2}\log_2\frac{1-pI}{I-pI}\right)\cdot I \tag{9.20}$$

经整理可解得

$$\begin{cases}\tilde{I}(p)=[p+(1-p)\cdot 2^{(2-4p)}]^{-1}\\ \tilde{P}(p)=p\cdot\tilde{I}(p)\end{cases} \tag{9.21}$$

(9.21)式中写成 $\tilde{I}(p)$ 与 $\tilde{P}(p)$ 纯粹是因为 $\tilde{I}$ 确实是 p 的函数。(9.21)式所决定的 $(\tilde{P}(p),\tilde{I}(p))$ 是唯一不亏不盈的协议所具有的形式。如果甲购买了保险单 $(\tilde{P}(p),\tilde{I}(p))$，万一他被解雇，他不去申请索取失业保险补偿而去寻找新工作的最优努力水平为

$$\begin{aligned}H^*(\tilde{P}(p),\tilde{I}(p))&=\left[\frac{1}{2}\log_2\left(\frac{1-\tilde{P}(p)}{\tilde{I}(p)-\tilde{P}(p)}\right)\right]^2\\&=\left[\frac{1}{2}\log_2\left(\frac{\tilde{I}(p)^{-1}-p}{1-p}\right)\right]^2\\&=\left[\frac{1}{2}\log_2(2^{2-4p})\right]^2\\&=(1-2p)^2\end{aligned} \tag{9.22}$$

对于不同的 p，我们已经得到了保险单 $(\tilde{P}(p),\tilde{I}(p))$，它从保险统计角度是公正的，相应于每一对 $(\tilde{P},\tilde{I})$，甲寻找新工作的最优努力水平 H^* 也由(9.22)式给出。问题仍然在于甲究竟应买多少保险，事实上，我们还有待于确定 p，从而得出甲花费 $\tilde{P}(p)$ 买失业保险，该保险单将保证他可以获得 $\tilde{I}(p)$ 的补偿。前面的讨论业已指出，倘若不存在道德风险，那么甲应当购买全额保险，在那里，$p=0.5$，$\tilde{I}(p)=1$，$\tilde{P}(p)=0.5$，$H^*(\tilde{P},\tilde{I})=0$。现在我们的问题是失业

保险博弈出现道德风险，这种应购买全额保险的讲法是否还是正确的呢？

甲购买了保险单$(\tilde{P}(p),\tilde{I}(p))$后将面临三种可能结局：他没有被解雇；他被解雇后又找到新工作；以及他被解雇但找不到新工作。这三个事件发生的概率和发生其中之一时某甲的效用可以用表 9.1 列出。

表 9.1

情况	发生概率	甲的效用
没被解雇	0.5	$1-\tilde{P}(p)=1-p\tilde{I}(p)$
被解雇但找到新工作	$0.5\pi(H^*(p,\tilde{P}(p),\tilde{I}(p))$ $=0.5\cdot\frac{1}{2}\log_2(\frac{\tilde{I}(p)^{-1}-p}{1-p})$ $=0.5-p$	$1-\tilde{P}(p)-H^*(\tilde{P}(p),\tilde{I}(p))$ $=1-p\tilde{I}(p)-(1-2p)^2$
被解雇但找不到新工作	$0.5\cdot[1-\pi(H^*(p,\tilde{P}(p),\tilde{I}(p))]$ $=0.5\cdot[1-\frac{1}{2}\log_2(\frac{\tilde{I}(p)^{-1}-p}{1-p})]$ $=p$	$\tilde{I}(p)-\tilde{P}(p)-H^*(\tilde{P}(p),\tilde{I}(p))$ $=(1-p)\tilde{I}(p)-(1-2p)^2$

甲选取一个使他的期望效用极大化的且保险公司不亏不盈的协议（这就是子博弈完美均衡所要求的第二点）：

$$\begin{aligned}&EU(\tilde{P}(p),\tilde{I}(p),H^*(\tilde{P}(p),\tilde{I}(p))\\&=0.5(1-p\cdot\tilde{I}(p))+(0.5-p)\cdot[1-p\cdot\tilde{I}(p)-(1-2p)^2]\\&\quad+p\cdot[\tilde{I}(p)-p\cdot\tilde{I}(p)-(1-2p)^2]\\&=0.5+p-2p^2\end{aligned}\tag{9.23}$$

这是一个关于 p 的二次型，它具有唯一的极大值点 $p^*=0.25$（就是说，购买的保险单规定，每补偿 100 元，需支付 25 元保险费）。以这个 p^*，某甲购买的最优保险单为

$$I^*=\tilde{I}(p^*)=1/[0.25+0.75\cdot2^{(2-4\times0.25)}]=\frac{4}{7}\approx0.57\tag{9.24}$$

$$p^*=\tilde{P}(p^*)=0.25\times0.57=0.14(\text{千元})\tag{9.25}$$

购买这样的保险单，万一他被解雇，他将花费如下代价以寻求新工作：

$$H^{*}=H^{*}(\tilde{P}(p^{*}),\tilde{I}(p^{*}))=(1-2\times0.25)^{2}=0.25 \quad (9.26)$$

而它找到新工作的概率为

$$\pi(0.25)=\sqrt{0.25}=0.50 \quad (9.27)$$

不难发现，甲购买的最优保险单不是全保险，只有这样的 57%的补偿率，使得甲在一旦失业后会主动地努力寻找新工作且找到新工作的概率为 50%，他将发现自己有 25%(0.5×0.5)时间失业，并且必须忍受 430 元的收入损失。

正是道德风险的存在，使得甲购买不完全保险。他这样做是为了得到较低保险价格的好处，因为在没有道德风险的情况下，我们看到公正的价格是 $p=P/I=0.5$，而在道德风险出现的失业保险市场，其价格为 $p^{*}=0.25$。对于这些结果，我们不能马上清楚由于道德风险的问题，甲是否在买保险博弈中受到了损害。现在我们将通过讨论指出，一般情况下，甲的确受到了损害。

假设保险公司可以在无需成本的情况下观察到甲的寻求工作的努力程度 H。此时，保险公司可以制定价格 p，在理论上它依赖于 H 与 I。但是由于只有 H 影响到失业风险，因此不再有任何理由使 p 依赖于 I。现在，委托代理博弈发生了变化，仍然是保险公司首先行动并公布保险单一览表，但是价格表仅依赖于保险购买者甲的关于 H 的承诺而不依赖于 I。甲仍然是第二行动者，他选择 I——失业补偿水平。可是现在他必须预先承诺，一旦失业他寻求新工作的努力为 H。如果甲失去了工作而不像他原先承诺的那样努力寻求新工作，这一点被保险公司完全观察到(这是本次模型的假设)，由于甲违背了协议，保险公司不必为他支付任何东西。可以想象这是一个苛刻的待遇，可是倘若没有这样厉害的一着，甲为了获得低的保险价格，他将被激励而承诺高 H 值。然而当他真的置身于劳动力市场时，他则不会那么卖力地寻求新工作，保险公司

将发现自己按要求付出的高于从保险单中获得的收入。

继续假定保险公司仅提供保险统计公正的协议，那么均衡保险费一览表 $\tilde{\tilde{P}}(H)$ 将满足等式：

$$\tilde{\tilde{P}}(H)=0.5(1-\pi(H))\cdot I=0.5(1-\sqrt{H})I \qquad (9.28)$$

由于 p 不再依赖于 I，因此甲可以随自己的愿望来购买保险单 (P,I)。例如，他可以承担 H 的水平满足 $p=0.25$ 这样的义务，以这个价格，他可以购买 570 元的补偿保险单而付出保险费 140 元。这组数据恰好是上面存在道德风险时的最佳方案。简言之，甲把自己置身于存在道德风险的环境。然而，"他可以做的"与"他将要做的"是两码事。因为所提供的一切保险都是公正的，在完全没有道德风险的情况下，无论甲选择 H 如何，他应当购买全保险。就是说，如果他承诺 H 的水平使得 $p=0.25$，那么在没有道德风险的情况下他购买的保险单中的补偿额不应当是 570 元而应当是 1 000元。甲可以选择存在道德风险时的最佳的 H 与 I，可是他不必这样去做。这表示当保险公司可以监控他的 H 从而略去了道德风险时，他的处境不会更糟。他可以有严格地更好一些的结果。观察一下如下的计算结果就可以对这段话有所理解。

甲应当选择最优的努力程度 H^* 使得自己的效用达到极大。为与前面的讨论可比较起见，仍取 $U(Y,H)=\log_2(Y)-H$。由于略去了道德风险，故应取 $I=1$，于是 H^* 应极大化

$$\log_2(1-\tilde{\tilde{P}}(H))-H=\log_2(0.5+0.5\sqrt{H})-H \qquad (9.29)$$

不妨记 $\sqrt{H}=x$，在 $\log_2(0.5+0.5x)-x^2$ 中对 x 求导数并使之为 0，得

$$\frac{1}{(1+x)\ln 2}-2x=0 \qquad (9.30)$$

解得 $x=\sqrt{H^*}\approx 0.485$，即 $\pi(H^*)=\sqrt{H^*}\approx 0.485$。因此在均衡状态，任何一个月中，甲有 $0.5\times(1-0.485)=0.2576\approx 0.26$ 的

机会处于失业状态。

§9.2 道德风险与非自愿失业

1. 模型

设某公司的转运仓库雇佣一批搬运工，他们将承担搬运、清扫、整理等各种各样杂活，正因为这个打杂的工作性质使得任务不断变更，给工人创造了偷懒或开小差的机会。工人不偷懒的时间比例记作 H，显然 $0 \leqslant H \leqslant 1$。$H$ 当然可以看作工人的努力程度。倘若有 L 个工人为公司干活，假设每个工人的努力水平均等于 H，那么 $L \cdot H = L_H$ 是实际工作的劳动力。公司的收益 R 依赖于 L_H 而不是依赖于 L。如同上节那样，我们可以作一个合理的假设：R 是 L_H 的递增函数，但递增的速度却是 L_H 的递减函数。并令 $R(0)=0$，表明劳动力对于公司的收益是必不可少的。满足这些假设条件的 R 不止一个，简单的实例是 $R(L_H)=\ln(1+L_H)$。为了这些收益，公司当然得花费成本，该成本依赖于签约的人数 L 以及签约工资 W。当然，公司愿意按实际工作的小时数 L_H 支付工资，但 L_H 是无法真正观察到的。于是公司的成本是 $W \cdot L$。定义公司的效率工资为 $W_H = W/H$。现在，我们可以给出公司的效用函数：

$$V(W,L,H)=\ln(1+L \cdot H)-W \cdot L \tag{9.31}$$

由于工人在实际工作中有一定耗费，因此每个工人的效用函数应是 W 与 H 的函数，不妨假设为如下形式：

$$U(W,H)=W\left(1-\frac{3}{8}H\right) \tag{9.32}$$

如果一个工人离开公司，他可得到(社会)保留效用$\overline{U}$。工人没有任何理由就业于报酬效用低于保留效用的其他公司。

对于那些“很经常”地被逮到偷懒的工人，公司通过解雇给予惩罚。公司通过随机检查手段发现偷懒者。由法律或规则确定怎

样算“很经常”。工人偷懒被常常逮住从而遭到解雇的概率是工作努力程度 H 的递减函数，假定这个函数 $\pi(H)$ 对所有工人都相同，不妨设其取形式

$$\pi(H)=1-H \tag{9.33}$$

公司与它的每一个工人之间可以建立具有道德风险的委托——代理博弈模型。之所以说该模型具有道德风险，是因为一旦公司与工人签署了雇佣协议之后，工人的是否偷懒并不能被公司完全观察到。在这个博弈模型中，公司首先行动，它选择准备雇佣的工人数 L 和付给每个工人的工资额 W。工人在其后行动，每一个工人（假定）同时独立地进行决策，是接受还是拒绝雇佣。由于我们假设工人都是“一样的”，因此在讨论中考虑 L 个工人中一个代表性人物，记作 A，以代表全体工人。如果 A 拒绝公司所提出的工资额，那么博弈至此结束，公司获取的收益为 0，而 A 将得到社会保留效用 $\bar{U}$。如果 A 接受了公司的工资额并为公司工作，那么博弈继续展开，由 A 进行自己的第二次选择，此时 A 选择的是他的工作努力程度 H。众所周知，该努力程度不为公司所观察到。公司关于 H 的所有信息仅通过经常性地随机检查来得到。检查的真实结果只有“老天”最清楚，因此我们称“自然”是本博弈模型的最后行动者，它确定了 A 是否经常被逮住偷懒从而遭到解雇。读者不难根据上述顺序画出模型的博弈树，不过请注意到，由于第一与第二次行动有许许多多种可行方案，我们常用一个箭头且在箭头上方或旁边注上可能的行动，比如，第三次由工人 A 行动，决定自己的工作努力程度 H，不少博弈论教科书上采用如图 9.1 所示方式表示。

注意到 A 的工作努力程度 H 不可能被公司直接观察到，公司不能够签署一个 A 的工资依赖于 H 的合同。一般的劳务合同至多规定若工人违背公司有关规章制度则立即解聘等等。因此公司的策略只是简单的向量 (W,L)，工人 A 的策略则要复杂得多。



图9.1

首先，他应决定是否接受公司的聘用，然后他确定自己的工作强度H，而H依赖于A所得到的工资，我们可以用$H(W)$表示这种依赖关系。

显然，博弈有三个可能结局：第一个结局是A拒绝公司的劳务合同从而结束博弈。A由此得到的保留效用$\bar{U}$依赖于他在公司之外的就业机会和社会失业保险，A与公司都可视$\bar{U}$为外生的(exogenous)。第二个可能结局是A接受了公司提出的工资且为公司干活，没有因为过分的偷懒而被解雇。在这个结局里，公司的盈利为$\ln(1+L\cdot H)-W\cdot L$，A（每个工人）的盈利为$W(1-\frac{3}{8}H)$。第三个结局是A接受了公司提供的工资额并为公司工作，但是由于经常性地被发现在偷懒而遭到解雇。此时，我们总是假定公司可以立即找到另外一个新工人替代A工作并领取同样的报酬，而且为方便见，我们仍认可这个另外的新工人与原来所有工人都是"什么都一样"，即W一样，H也一样。这种假设使得我们在计算公司的盈利时不会遇到任何困难，它将与第二种结局时一样，至于A，他得到的将仅仅是保留效用$\bar{U}$。

必须指出，在我们建立的模型中，A作为代表性工人，其收益以一个人来计算，而公司的盈利计算则是面对他雇用的L个工人来进行。因此A若拒绝W，则由于A是代表性人物，相当于L个工人全拒绝，但到最后，尽管L个工人都像A一样地选择工作努力程度H，由于公司的检查是随机的，因此A被逮到经常偷懒这件事不等于L个工人同时经常地被逮到在偷懒。希望读者在学习这个博弈模型时应考虑到这一点，这也就是为什么A的盈利总是

以一个人计算，公司的盈利总以 L 个工人来计算的道理。

利用后退归纳法来解博弈问题。尽管博弈的最后一个行动者是“自然”，但这仅仅表示一旦 A 选择了 H，由于外生的“自然”行动“随机地”使结局呈现不确定性。因此这里的后退归纳要从真正的最后一个局中人的行动开始后退。A 就是符合这一要求的。在 W 与 $\bar{U}$ 为给定的条件下，A 选择努力程度 H，必须使自己的期望效用达到极大化，由于经常性地被逮住偷懒的概率为 $\pi(H)$，因此，实际上，A 在选择 H，以使

$$EU(W,H)=\pi(H)\bar{U}+(1-\pi(H))W(1-\frac{3}{8}H) \qquad (9.34)$$

极大化。由于 $\pi(H)=1-H$，因而

$$EU(W,H)=(1-H)\bar{U}+W\cdot H(1-\frac{3}{8}H)$$

$$=\bar{U}+(W-\bar{U})H-\frac{3}{8}W\cdot H^2 \qquad (9.35)$$

(9.35)式是个二次型，似乎不难求得它的极大值点 H^*。但是它的系数直接影响到我们所需要的结论。例如当 $W\leqslant\bar{U}$ 时，H 与 H^2 前的系数均为非正，显然此时不如取 $H^*=0$ 为好。这很容易在直观上给予解释，当公司开出的工资不大于保留效用时，实际上工人 A 会选择“拒绝”，因此相当于选择了工作努力程度为 0。一般情况，我们可以得到一阶条件：

$$(W-\bar{U})-\frac{3}{4}WH=0 \qquad (9.36)$$

或

$$H=\frac{4}{3}\cdot\left(\frac{W-\bar{U}}{W}\right) \qquad (9.37)$$

我们知道，必应有 $H\leqslant1$，因而 $\frac{4}{3}\cdot\left(\frac{W-\bar{U}}{W}\right)\leqslant1$，为使(9.37)式合理，我们得到 W 与 $\bar{U}$ 的另一个关系式应为 $W\leqslant4\bar{U}$，当 $W>4\bar{U}$ 时，工作努力程度至多取 1。于是

$$H^*(W)=\begin{cases}0 & 若\ W\leqslant\bar{U}\\ \frac{4}{3}\left(\frac{W-\bar{U}}{W}\right) & 若\ \bar{U}<W\leqslant 4\bar{U}\\ 1 & 若\ 4\bar{U}<W\end{cases} \tag{9.38}$$

已经指出，当且仅当 $W\geqslant\bar{U}$ 时工人 A 才会接受公司所开出的工资，此时

$$EU(W,H^*)=\bar{U}+(W-\bar{U})H^*-\frac{3}{8}W(H^*)^2$$

$$\geqslant U+\frac{2}{3}\cdot\frac{(W-\bar{U})^2}{W}\geqslant\bar{U} \tag{9.39}$$

即，当且仅当 $EU(W,H^*)\geqslant\bar{U}$ 时，A 会接受公司的开价。现在往后退，考虑公司的有关策略，由(9.38)式，公司不会宣布 $W<\bar{U}$，因为这将导致招不到工人，同时公司也不会令 $W=\bar{U}$，因为这将引起工人 A 即使接受，其工作努力程度为 $H^*(\bar{U})=0$。当然，(9.38)式也告诉我们，公司若取 $W>4\bar{U}$(高工资)，A 所选择的工作努力程度至多取 1，任何公司不会傻到这种地步！用博弈论的术语来说，任何小于 $\bar{U}$ 或大于 $4\bar{U}$ 的 W 都是公司的劣策略。在工人 A 选定 $H^*(W)$情况下，考虑公司的效用函数

$$V(W,L)=\ln(1+L\cdot H^*(W)-W\cdot L) \tag{9.40}$$

这里 $V(W,L)$是 W 与 L 的二元函数，为使 V 达到极大化，其相应的一阶条件为

$$\frac{\partial V(W\cdot L)}{\partial W}=0=\frac{\partial V(W\cdot L)}{\partial L} \tag{9.41}$$

(9.41)式似乎很有理，但由于 L 一般取正整数，不是连续统的，因此对 L 求偏导数似乎令经济管理类读者无法接受，其实，我们可以将 L“视作”连续统变量，在求得 L^* 之后，再将它合理地“整数”化。在 W 为非劣策略情况下，将 $H^*(W)=\frac{4}{3}\cdot\frac{(W-U)}{W}$ 代入(9.40)式，得

$$V(W,L)=\ln\left(1+\frac{4}{3}\cdot\frac{L(W-\bar{U})}{W}\right)-L\cdot W \qquad (9.42)$$

因此

$$\frac{\partial V}{\partial W}=\frac{4L\bar{U}}{W[3W+4L(W-\bar{U})]}-L \qquad (9.43)$$

$$\frac{\partial V}{\partial L}=\frac{4(W-\bar{U})}{3W+4L(W+\bar{U})}-W \qquad (9.44)$$

利用(9.41)式，化(9.43)式与(9.44)式为

$$\begin{cases}4\bar{U}=3W^2+4LW(W-\bar{U})\\4(W-\bar{U})=3W^2-4LW(W-\bar{U})\end{cases} \qquad (9.45)$$

由(9.45)式，必有 $4\bar{U}=4(W-\bar{U})$，即

$$W^*=2\bar{U} \qquad (9.46)$$

以 $W^*=2\bar{U}$ 代回(9.45)第一式，得 $L=\frac{1-3\bar{U}}{2\bar{U}}$，由于 L 为非负，故可得

$$L^*=\begin{cases}\frac{1-3\bar{U}}{2\bar{U}} & 若\ \bar{U}<\frac{1}{3}\\0 & 若\ \bar{U}\geqslant\frac{1}{3}\end{cases} \qquad (9.47)$$

此时

$$H^*(W^*)=\frac{4}{3}\cdot\frac{W^*-\bar{U}}{W^*}=\frac{2}{3} \qquad (9.48)$$

在 $\bar{U}<\frac{1}{3}$ 时，博弈模型的子博弈完美均衡为：公司招收 $\frac{1-3\bar{U}}{2\bar{U}}$ 个工人(实际上应考虑最邻近的整数值)，付给每个工人的工资为 $2\bar{U}$，工人接受并为公司工作，同时选择他的工作努力程度为 $\frac{2}{3}$。

在子博弈完美均衡中，每个工人有三分之一的时间在偷懒或开小差，这是一个比较高的数且很令人不高兴，其原因在于在监控工人方面公司仅有有限的能力。如果公司希望得到较高的工作努力程度，那么它只有付出较高一些的工资。

2. 模型的趋于完善：非自愿失业(involuntary unemployment)的必然性

我们在前面得到的均衡在一个很重要的方面是有欠缺的：我们将工人的保留效用 $\bar{U}$ 看作是外生的。在实际世界，这个量由劳动力市场确定。本小段里，我们将通过试图得出 $\bar{U}$ 以使博弈模型趋于完善。工人 A 除了为公司干活还有什么其他的内容呢？如果用统计学的语言来讲，原假设为"A 替公司打工"那么备择假设是什么呢?在趋于完善的模型中，我们认定"A 要么为其他公司打工，要么失业"。假设有若干"一模一样"的公司进行竞争，为使模型尽可能地简单一些，所谓一模一样，表示这些公司从事同样的业务，需要同样的工人，具有同样的效用函数以及拥有同样的监控工人是否偷懒的手段。同时我们假定，这些公司同时且独立地作出有关工资与雇工决策。

先计算一下，在上述子博弈完美均衡中，任何一个雇工 A 的期望效用应为

$$EU(W^*,H(W^*))=\frac{1}{3}\bar{U}+\frac{2}{3}(2\bar{U})(1-\frac{3}{8}\cdot\frac{2}{3})=\frac{4}{3}U \tag{9.49}$$

或者说，为任何一个其他公司工作的工人在均衡中的期望效用均为 $\frac{4}{3}\bar{U}$。而且，在均衡状态，所有公司支付给工人的最"合适"工资都是 $W^*=2\bar{U}$。假如这个工资是市场出清工资，那么任何一个被某公司辞退的工人立刻可以在另一家公司找到同样的工作，获得同样的工资。倘若情况的确如此，那么工人总是受雇且保留效用 U 必定等于在其他公司受雇时的子博弈完美期望效用，这个期望效用等于 $\frac{4}{3}U$。只要 $U>0$ 的话，矛盾自然而然地呈现在大家面前，因为这个 $\frac{4}{3}\bar{U}$ 是在保留效用 U 给定情况下的均衡状态中计算得来

的。看来，市场出清工资与子博弈完美均衡存在着相当的不协调，或者说市场出清工资不可能是均衡工资，除非市场上存在着失业——非自愿失业。

均衡中失业的存在蕴含着每个公司面对劳动力的过分供应。劳动力市场供大于求，分配工作的一个办法是采用先到者先安排。这是根据工人都是“一模一样”这个假设而合理制定的，那些刚被解雇的工人回到了失业线，而那些站在失业线前列的工人则去顶替他们的工作。失业的“魅力”在于可以对那些“不老实”的雇员给予解雇威胁，同时提供了更大效果的动力。设 L_T 表示工人的总数，为简便起见，我们假定 L_T 是一个固定数。倘若有 n 家公司，每家公司雇佣 L^* 个工人，那么失业人数等于 L_T-nL^*。我们已知每一家公司将逮住三分之一经常性偷懒的工人并解雇他们。公司将立即解决同样数目的失业人员的就业问题。如果用记号 T_u 表示，一旦工人被解雇，他们预期再次被聘的周期数，该周期数显然等于

$$T_u=\frac{\text{失业工人的存量}}{\text{每期新雇用工人的流量}}=\frac{L_T-nL^*}{\frac{nL^*}{3}}=\frac{3L_T}{nL^*}-3 \qquad (9.50)$$

由于 $L^*=\frac{1-3\overline{U}}{2\overline{U}}$，代人(9.50)式得

$$T_u=\frac{6\overline{U}(L_T+\frac{3}{2}n-\frac{n}{2\overline{U}})}{n(1-3\overline{U})} \qquad (9.51)$$

举例来说，若有 100 名像上节中 A 一样的工人，两个公司雇佣这类工人，每家雇 30 人，在均衡状态中，任何时候有 40 名失业工人。对于新的刚被解雇的工人，排到失业队伍的前列并再被雇佣需要两个周期。(9.51)式强烈地暗示，周期数 T_u 与保留效用 $\overline{U}$ 有着密切的关系，考虑 $\overline{U}\neq\frac{1}{3}$，在(9.51)式中对 $\overline{U}$ 求导，有

$$\frac{\mathrm{d}T_u}{\mathrm{d}\,\overline{U}}=\frac{6L_T}{n(1-3\overline{U})^2}>0 \qquad (9.52)$$

当保留效用 $\bar{U}$ 增大时，刚被解雇的工人为再就业所需周期数也随之增大。这一点很符合实践。

这个博弈模型清晰地告诉我们，由于道德风险的存在，必然会产生非自愿失业。

第四部分　不完全信息静态博弈

前面两部分，我们研究了完全信息博弈，其主要特点之一是“局中人的盈利是共同知识”，在经济学的许多应用中，这只不过是理想的模式，通常，至少有一个局中人不知道（或不确定）其他局中人的盈利函数，于是信息便成为不完全的。从现在开始，我们将对不完全信息博弈进行研究，此类博弈也称作 Bayes 博弈。在第三部分中，主要研究不完全信息静态博弈，最普通的例子是密封投标拍卖（sealed-bid auction）：关于待售的物品，每一个投标者都有自己的估价，但却不知道任何其他投标者的估价，投标方式是通过将标价放在密封信封中递交，这样局中人的行动可以认为是同时的，也就是符合静态博弈的要求。

静态博弈既然具有不完全信息，那么对均衡必然应有新的理解，这是第十章所要解决的问题。此外，以拍卖为例，如果将游戏规则作一定改变，势必引起博弈预测结局的变化，这涉及到博弈的机制设计，该方面内容将在第十一章中详细介绍。

第十章　Bayes 博弈与 Bayes 均衡

静态 Bayes 博弈以及相应的 Bayes Nash 均衡(今后将简称 Bayes 均衡,就像前面常用均衡替代 Nash 均衡一样)的定义比较抽象且复杂,我们将以一个简单的例子——不对称信息下 Cournot 竞争以引进主要思想。

§10.1　非对称信息下的 Cournot 竞争

回想 Cournot 寡头模型,需求函数为 $P(Q)=a-Q=a-(q_1+q_2)$,公司 1 的成本函数是 $C_1(Q)=cq_1$,而公司 2(假设它是产业的新参加者,或者,也可能刚发明了一种新技术)的成本函数是

$$C_2(q_2)=\begin{cases} C_H q_2 & \text{概率为 } \theta \\ C_L q_2 & \text{概率为 } 1-\theta \end{cases} \qquad \text{其中 } C_L<C_H$$

公司 2 知道自己成本函数和公司 1 的成本函数,但是公司 1 除了知道自己的成本函数之外,对公司 2 的成本函数无法完全确定,它仅仅知道公司 2 的边际成本以概率 θ 为 C_H 而以概率$(1-\theta)$为 C_L,很自然地,我们称该模型中(二个局中人)的信息是非对称的。上述非对称信息是共同知识:公司 1 知道公司 2 有比自己优越的信息,公司 2 知道公司 1 是知晓这一点的,等等。

毫无疑问,公司 2 对于自己不同的边际成本应选择不同的产量,而公司 1 则必须预测公司 2 可能选择什么样的边际成本,然后再依据自己的边际成本来选择自己的产量以极大化自己的盈利。

具体地说，若以 $q_2^*(C_H)$ 与 $q_2^*(C_L)$ 分别表示公司 2 对于不同边际成本所选择的产量，公司 1 的产量选择 q_1^* 却只有一个，因为它的边际成本为唯一。如果公司 2 取高成本 C_H，对于公司 2，它将选择 $q_2^*(C_H)$ 以满足

$$\max_{q_2}[(a-q_1^*-q_2)-C_H]\cdot q_2 \tag{10.1}$$

同样地，如果公司 2 为低成本，它将选择满足下式的 $q_2^*(C_L)$

$$\max_{q_2}[(a-q_1^*-q_2)-C_L]\cdot q_2 \tag{10.2}$$

至于公司 1 由于它只知道公司 2 取 C_H 与 C_L 的二点分布，因此，它选择的 q_1^* 应当使如下期望盈利达到极大化：

$$\max_{q_1}\{\theta[(a-q_1-q_2^*(C_H))-c]q_1 + (1-\theta)[(a-q_1-q_2^*(C_L))-c]q_1\} \tag{10.3}$$

这三个最优化问题所得到的首阶条件是

$$q_2^*(C_H)=(a-q_1^*-C_H)/2 \tag{10.4}$$

$$q_2^*(C_L)=(a-q_1^*-C_L)/2 \tag{10.5}$$

$$q_1^*=\{\theta[a-q_2^*(C_H)-c]+(1-\theta)[a-q_2^*(C_L)-c]\}/2 \tag{10.6}$$

综合(10.4)式～(10.6)式，不难解得

$$q_2^*(C_H)=\frac{a-2C_H+c}{3}+\frac{1-\theta}{6}(C_H-C_L) \tag{10.7}$$

$$q_2^*(C_L)=\frac{a-2C_L+c}{3}-\frac{\theta}{6}(C_H-C_L) \tag{10.8}$$

$$q_1^*=\frac{a-2c+\theta C_H+(1-\theta)C_L}{3} \tag{10.9}$$

如果 $C_H=C_L=C_2$，那么得到的恰为信息 Cournot 竞争的 Nash 均衡解。可见，不完全信息下，Cournot 竞争的预测解与完全信息静态 Nash 均衡解之间的确存在着差异。在不完全信息情况，我们有

$$q_2^*(C_L)-q_2^*(C_H)=\frac{1}{2}(C_H-C_L)>0 \tag{10.10}$$

(10.10)式表明，当边际成本较高时，公司 2 将少生产一些；反之，当边际成本较低时，它将略多生产一些，非常符合常理。然而，(10.7)式与(10.8)式又蕴含了

$$\frac{(a-2C_H+c)}{3}<q_2^*(C_H)<q_2^*(C_L)<\frac{(a-2C_L+c)}{3} \quad (10.11)$$

我们仅需讨论最左边一个不等式，倘若公司 2 的成本为 C_H 且该信息为公司 1 所完全掌握，那么问题又回到完全信息静态博弈，公司 2 的均衡策略行动应取 $q_2^*=(a-2C_H+c)/3$，在不完全信息静态博弈中，之所以 $q_2^*(C_H)$ 大于 $(a-2C_H+c)/3$，正是因为公司 2 考虑到公司 1 未必正确地预测到自己的边际成本究竟是高还是低，从而针对这一事实公司 2 作出相应的反应。

在非对称信息 Cournot 竞争例子中，显然信息是不完全的，我们考虑的是纯策略均衡解。公司的策略显然是产量选择 q_1 与 q_2，似乎与完全信息静态博弈一样，但是注意到公司 2 的盈利函数可以表示为 $u_2(q_1,q_2,t_2)$，其中 t_2 等于 C_H 或 C_L，我们称 t_2 为公司 2 的类型，类型空间 $T_2=\{C_H,C_L\}$。已经知道，每一个类型 t_2 对应于局中人 2 可能有的不同的盈利函数。

用一般的话来描述，假如 n 个局中人进行静态(即同时行动)博弈，局中人 i 的行动空间是 $A_i(i=1,2,\cdots,n)$，他有数个(不妨设为 2 个)可能的盈利函数。则称局中人 i 有两个类型 t_{i1} 与 t_{i2}，他的类型空间 $T_i=\{t_{i1},t_{i2}\}$，局中人 i 的两个盈利函数为 $u_i(a_1,\cdots,a_n;t_{i1})$ 与 $u_i(a_1,\cdots,a_n;t_{i2})$。因此我们可以利用每一个局中人的类型对应于该局中人可能有的不同的盈利函数这样一个想法，以表示局中人有不同的可行行动集的可能性。例如局中人 i 的可行行动集如下：以概率 p 为 $\{a,b\}$，以概率 $(1-p)$ 为 $\{a,b,c\}$，那么我们可以说局中人 i 有两个类型 t_{i1} 与 t_{i2}，其中 t_{i1} 的概率为 p，t_{i2} 概率为 $(1-p)$，然后我们定义这两种类型时局中人 i 可行行动集是 $\{a,b,c\}$，只不过对类型 t_{i1} 时规定局中人 i 取行动 c 的盈利是 $-\infty$(因此

他肯定不会取 c)。在非对称信息 Cournot 竞争中，公司 2 有两个盈利函数，因为他有两种类型 C_H 与 C_L，而公司 1 由于只有一种类型 c，从而只有一个可能的盈利函数，即 $u_1=[(a-q_1-q_2)-c]q_1$。此时 $T_1=\{c\}$，$T_2=\{C_H,C_L\}$。

引进了局中人 i 的类型这一概念，于是我们称局中人 i 知道自己的盈利函数就等价于称局中人 i 知道自己的类型。同样地，称局中人 i 关于其他局中人的盈利函数可能无法吃准等价于局中人 i 关于其他局中人的类型，$t_{-i}=(t_1,\cdots,t_{i-1},t_{i+1},\cdots t_n)$ 究竟是什么的判断可能是不正确的。习惯上用 T_{-i} 来表示所有可能的 t_{-i} 值的全体。而且我们使用概率分布 $p_i(t_{-i}|t_i)$ 表示局中人 i 在给定自己的类型知识的条件下关于其他局中人类型 t_{-i} 的信念(belief)。在本章讨论以及应用分析中，假定局中人的类型分布是独立的，于是信念 $p_i(t_{-i}|t_i)$ 不依赖于 t_i，则简记局中人 i 的信念为 $p_i(t_{-i})$。

将类型与信念这两个新概念与完全信息静态博弈正则型表示中所熟悉的元素相结合，则产生了静态 Bayes 博弈的正则型表示：

定义 10.1 n 人静态 Bayes 博弈的正则型表示确定了局中人的行动空间 $A_1,A_2,\cdots,A_n$，和他们的类型空间 $T_1,\cdots,T_n$，以及他们的信念 $p_1,\cdots,p_n$，还有各局中人的盈利函数 $u_1,u_2,\cdots,u_n$。局中人 i 的类型 t_i 为局中人 i 私人所知，它确定了局中人 i 的盈利函数 $u_i(a_1,\cdots,a_n;t_i)$，$t_i\in T_i(i=1,2,\cdots,n)$。局中人 i 的信念 $p_i(t_{-i}|t_i)$ 描述了局中人 i 在给定自己的类型 t_i 的条件下关于其他 $(n-1)$ 个局中人的可能类型 t_{-i} 的不确定性。我们将这个博弈表示为 $G=\{A_1,\cdots,A_n;T_1,\cdots,T_n;p_1,\cdots,p_n;u_1,\cdots,u_n\}$。

§10.2 Harsanyi 转换

信息的不完全使得对博弈的分析变得更复杂，因为信息不完

全的局中人必须预测其他局中人的类型。对此，Harsanyi 提出了一个建模的方法，引进由“自然”(nature)先期采取行动，它确定了(某)局中人的类型。这样，局中人 i 关于局中人 j 的类型的不完全信息就可以转换成关于自然行动的不完美信息，于是我们可以使用以前熟知的标准手段(或技巧)来分析新的转换博弈。为了详尽地解释 Harsanyi 转换(The Harsanyi Transformation)，让我们从最简单的进入—阻扰博弈(entry-deterrence game)着手。

两个局中人：在位者(incumbent)公司和可能进入者(entrant)。在位者必须决定是否扩展其工厂设备的容量，同时，进入者必须决定是否进人市场并与在位者竞争还是呆在市场之外。增加设备容量使在位者以较低成本生产较高产量。正因为此，当且仅当在位者不扩展时，“进人”对进入者是有利的。但是，在位者的盈利依赖于两个事实：进入者是否进入和扩展设备的成本。成本可能高也可能低。相应于高低成本时的盈利矩阵分别列于图 10.1 中。

		在位者（扩展成本低）扩展	在位者（扩展成本低）不扩展	在位者（扩展成本高）扩展	在位者（扩展成本高）不扩展
进入者	进入	−2,2	1,1	−1,−1	1,1
	不进入	0,4	0,3	0,0	0,3

扩展成本低　　扩展成本高

图 10.1

扩展成本的高或低，是在位者的两个类型，其信息属在位者私人所有，进入者并不知道，于是进入者关于进行中的博弈具有不完全信息。由图 10.1 可知，这相当于有两个可能的静态博弈，究竟是左边的还是右边的，这是需要进入者预测的。

然而，不管将进行的是哪一个博弈，从进入者的观点，当且仅当在位者不扩展的时候它希望进人。而从盈利矩阵可以看到，如果扩展成本低，在位者具有一个占优策略：扩展。而当扩展成本高时，“不扩展”则是他的占优策略。根据这些分析，扩展成本低，进入者

应采取“不进入”策略，而如果扩展成本高，他应当取“进入”。只要进入者能够正确地预测在位者的类型，它就能正确地采用有利于自己的行动。否则，它将处于“左右为难”地步。

因此，进入者必须对图 10.1 中究竟是左面还是右面是将要进行的博弈这一点形成信念，就像他必须对在位者究竟取“扩展”还是“不扩展”形成信念一样。基于这个事实，Harsanyi 提出将进入—阻扰博弈建模为一个新的博弈，那里有着三个局中人：自然、进入者与在位者。在位者可以是两种类型：高成本或低成本。现在，自然首先行动，由它来确定在位者到底属于何种类型。一旦自然确定了在位者的类型，在位者本身对此完全了解，从而他也会知道自己的盈利函数，但是进入者无法知道这些事情。面进入者由于只有一个类型，因此他的盈利函数显然是博弈双方的共同知识。自然选取在位者类型为高成本的概率称之为“高成本”类型的先验信念(prior belief)，假定这也是共同知识，不妨设为 p。我们对这个新博弈的一个解释是在位者与进入者由自然进行随机配对，高成本在位者配进入者的可能性为 p。一旦自然完成配对工作，事实上也就确定了图 10.1 中哪一个博弈将会进行。新博弈中规定自然先行动，然后在位者与进入者再同时行动，我们可以用博弈树来描述这个过程，为方便起见，记自然为 N，在位者为局中人 2，进入者为局中人 1。图 10.2 中自然选择在位者类型的概率用方括号里的数表示。读者对这个图至少当有似曾相识的感觉，因为这个博弈树描述了完全但不完美信息的动态博弈。所谓 Harsanyi 转换实际上就是巧妙地引入了“第三者”——自然，从而将复杂的不完全信息博弈转换为我们熟知的完全但不完美信息的博弈。不难设想，我们所要研究的 Bayes 博弈中的 Nash 均衡(有时称作 Harsanyi Bayes 均衡，经常地，干脆简单地叫作 Bayes 均衡)精确地就是转换了的完全但不完美信息博弈中的 Nash 均衡。

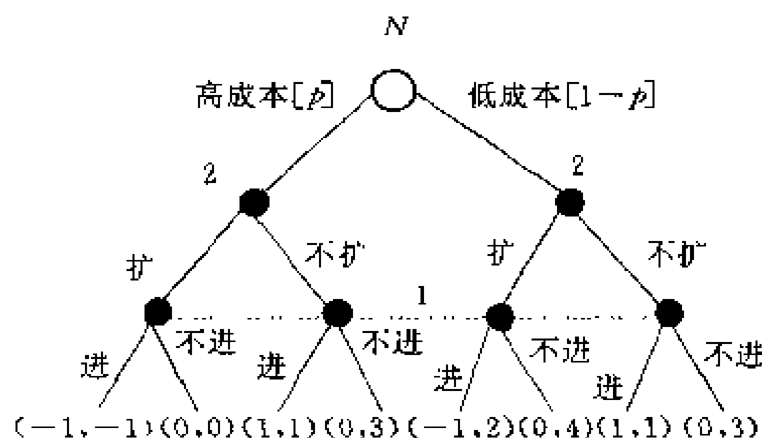


图10.2

继续以进入—阻扰模型为例，为具体化，令 $p=\frac{1}{3}$。在转换博弈中，进入者的策略仍然是要么进入要么不进；在位者的策略则称为"行动对"，"对"中第一个元素表示高成本时所取的行动，"对"中第二个元素表示低成本时所取的行动，例如(不扩展，扩展)表示当成本为高时在位者采取不扩展而当成本低时在位者采取扩展行动。如果策略剖面为{进入，(不扩展，扩展)}，花括号中第一个元素指进入者取"进入"策略，后面括号的意思如刚才所述。那么，以概率$\frac{1}{3}$，进入者的盈利为1，而以概率$\frac{2}{3}$，他的盈利为－1。因此进入者的期望盈利为$\frac{1}{3}\times 1+\frac{2}{3}\times(-1)=-\frac{1}{3}$。如果该策略剖面中进入者取行动"不进入"，显然他的期望盈利为0。由于(不扩展，扩展)是在位者的严优策略，不难看出{不进入，(不扩展，扩展)}是博弈的Nash均衡。当采用这个均衡策略时，有两个可能结局分别对应于在位者不同的类型。这个均衡与我们前面所说的"如果扩展成本低，应取(不进入，扩展)，而如果成本高时，则应取(进入，不扩展)"有着差异，差异存在于当成本高时的情况。为什么会造成差异呢？因为引号中概述的策略剖面指出了在进入者采取行动之前，他无法知晓自然究竟赋予在位者何种类型。可见，{不进入，(不扩展，

扩展)}作为新博弈的 Nash 均衡无疑是安全且保险的。

§10.3 Bayes 均衡

根据上一节 Harsanyi 转换，我们对 Bayes 均衡至少有了初步的了解，本节将对 Bayes 均衡作正式定义。

从进入—阻扰模型的解来看，我们关心的是比较局中人期望盈利的大小，设定 n 人静态 Bayes 博弈为 $G\{A_1,\cdots,A_n;T_1,\cdots,T_n;p_1,\cdots,p_n;u_1,\cdots,u_n\}$，局中人 i 的盈利函数不仅依赖于行动剖面 $(a_1,\cdots,a_n)$ 而且依赖于所有的类型 $(t_1,\cdots,t_n)$，故可记为 $U_i(a_1,\cdots,a_n;t_1,\cdots,t_n)$。为求期望盈利，需要计算信念 $p_i(t_{-i}|t_i)$。设自然按照先验分布 $p(t)$ 抽取类型向量 $t=(t_1,\cdots,t_n)$，这是一个共同知识，当自然向局中人 i 展示其类型 t_i 时，局中人 i 可以通过 Bayes 法则计算信念 $p_i(t_{-i}|t_i)$—— 因为这是一个条件概率。

$$p_i(t_{-i}|t_i)=\frac{p(t_{-i},t_i)}{p(t_i)}=\frac{p(t_{-i},t_i)}{\sum_{t_{-i}\in T_{-i}}p(t_{-i},t_i)} \qquad (10.12)$$

如前所述，我们常假设局中人的类型是随机独立的，于是信念 $p_i(t_{-i}|t_i)$ 将不依赖于 t_i，即为 $p_i(t_{-i})$，计算公式如下：

$$\begin{aligned}p_i(t_{-i})&=p(t_1,\cdots,t_{i-1},t_{i+1},\cdots,t_n)\\&=\sum_{t_i\in T_i}p(t_1,\cdots,t_{i-1},t_i,t_{i+1},\cdots,t_n) \qquad (10.13)\end{aligned}$$

局中人 i 的一个策略是类型 t_i 的函数 $s_i(t_i)$，即对类型空间 T_i 中的每一个类型 t_i，$s_i(t_i)$ 确定了在自然抽取类型 t_i 时局中人 i 从可行集 A_i 所选择的行动。当所有的局中人采取了策略剖面 $S=\{s_1(t_1),\cdots,s_n(t_n)\}$ 时，类型 t_i 的局中人 i 的条件期望效用(conditional expected utility，简记为 CEU)可计算为

$$EU_i(s,t_i)=\sum_{t_{-i}\in T_{-i}}U_i(s_1(t_1),\cdots,s_{i-1}(t_{i-1}),s_i(t_i),s_{i+1}(t_{i-1}),\cdots,$$

$$s_n(t_n), t_1, \cdots, t_{i-1}, t_i, t_{i+1}, \cdots, t_n) p_i(t_{-i}|t_i)$$

$$= \sum_{t_{-i} \in T_{-i}} U_i(s_{-i}(t_{-i}), s_i(t_i), t_i, t_{-i}) \cdot p_i(t_{-i}|t_i) \quad (10.14)$$

现在我们来定义 Bayes 均衡，不管定义中记号的复杂性，Bayes 均衡的中心思想既简单又为人们所熟悉：每一个局中人的策略行动必须是其他局中人策略行动的最佳反应。简单一句话，Bayes 均衡就是 Bayes 博弈中的 Nash 均衡。

定义 10.2 在静态 Bayes 博弈 $G\{A_1, \cdots, A_n; T_1, \cdots, T_n; p_1, \cdots, p_n; u_1, \cdots, u_n\}$ 中，策略 $s^* = (s_1^*, \cdots, s_n^*)$ 是一个（纯策略）Bayes 均衡，当且仅当，对每一个局中人 i 和 T_i 中的每一个类型 t_i 以及局中人 i 的每一个其他策略 $\hat{s}_i(t_i)$，成立

$$EU_i(s^*, t_i) \geqslant EU_i(s_1^*(t_1), \cdots, s_{i-1}^*(t_{i-1}), \hat{s}_i(t_i), s_{i+1}^*(t_{i-1}), \cdots, s_n^*(t_n), t_i) \quad (10.15)$$

就是说，无论局中人属何种类型，局中人的策略一定是关于其他局中人策略行动的最佳反应。

一个静态 Bayes 博弈，如果 n 为有限，$A_1, \cdots, A_n$ 以及 $T_1, \cdots, T_n$ 均为有限集合，那么称为有限静态 Bayes 博弈，它必定存在（至少一个）Bayes 均衡，当然也许它是一个混合策略。正因为 Bayes 均衡是 Bayes 博弈中的 Nash 均衡，因此有关 Bayes 均衡存在性的证明必将紧密地平行于完全信息有限博弈中混合策略 Nash 均衡存在性的证明，我们在这里就不再赘述。

从定义出发，我们再一次地核实{不进入，（不扩展，扩展）}的确是进入—阻扰模型的 Bayes 均衡。由 Harsanyi 转换，我们检查它的确是新的进入—阻扰模型（即完全但不完全信息博弈）的 Nash 均衡，为此我们建立新博弈的策略型式如图 10.3 所示。

在位者

		（扩展，扩展）	（不扩展，扩展）	（扩展，不扩展）	（不扩展，不扩展）
进入者	进入	(−1),(−1,2)	$(-\frac{1}{3})$,(1,2)	$\frac{1}{3}$,(−1,1)	1,(1,1)
	不进入	0,(0,4)	0,(3,4)	0,(0,3)	0,(3,3)

图 10.3

其中，每格的条件期望盈利为：(进入者，(高成本在位者，低成本在位者))。

现在我们将证实每一个局中人的均衡策略是关于其他局中人策略的最佳反应而不管局中人的类型究竟如何。在图 10.3 中，观察第一行，显然向量(1,2)优于(−1,2)，(−1,1)与(1,1)，而对于第二行，(3,4)优于(0,4)，(0,3)与(3,3)。从列的方向考虑，前二列，以“不进”时的盈利较大，而后二列，则“进入”的盈利($\frac{1}{3}$或 1)较大。于是我们立即得到了该盈利矩阵在“不管在位者类型究竟如何”的情况下的唯一 Nash 均衡：{不进入，(不扩展，扩展)}。

§ 10.4 Bayes 均衡的若干例子

本节我们将举一些 Bayes 均衡的例子，这些例子对我们进一步理解 Bayes 博弈与 Bayes 均衡有一定益处，况且有些例子在经济管理方面有一定的意义。

例 10.1 不完全信息下提供公共财产

公共财产的供给产生了著名的搭便车(free-rider)问题。当公共财产得到提供时，每一个人都将受益，可是每一个(理性)人又都希望由其他局中人承担供应公共财产的成本。现实中存在着大量公共财产范例，我们在这里仅考虑一个相当简单的例子：有两个局中人，$i=1,2$。他们同时决定是否向公共财产捐款，捐款可以用 0−1 决策来描述，即要么捐款，要么不捐款。如果他们两个人之间至少有一个

捐款，则每个局中人得益为1；如果没有一个人捐款给公共财产则大家均获益0。局中人 i 捐款额为 c_i。盈利矩阵如图10.4所示。

	捐款	不捐款
捐款	$1-c_1, 1-c_2$	$1-c_1, 1$
不捐款	$1, 1-c_2$	$0, 0$

图 10.4

只要有人捐款则各人获益1，这是共同知识，至于捐款多少，c_1 与 c_2 分别为局中人1与2的私人信息。但是，双方都相信"c_i 独立地来自 $[\underline{c}, \bar{c}]$ 上一个连续且严增累积分布函数 $P(\cdot)$"是共同知识，其中 $\underline{c}<1<\bar{c}$（因此 $P(\underline{c})=0, P(\bar{c})=1$ 为显然），这里与前面有所不同的是，局中人 i 的类型 c_i 来自连续的类型空间 $[\underline{c}, \bar{c}]$，$(i=1, 2)$。

由假设，捐款是0—1决策，因此局中人 i 的纯策略 $s_i(c_i)$ 是从 $[\underline{c}, \bar{c}]$ 到 $\{0,1\}$ 二点的一个函数，$s_i(c_i)=0$ 表示不捐款，$s_i(c_i)=1$ 表示类型 c_i 的局中人 i 采取捐款的策略行动。捐款额恰为类型 c_i。不难写出局中人 i 的盈利函数：

$$u_i(s_i, s_j, c_i)=\max(s_1, s_2)-c_i s_i \qquad (i\neq j) \qquad (10.16)$$

(10.16)式与图10.4完全一致，因为当两人都不捐款时，$s_1=s_2=0$，立得 $u_i(s_i, s_j, c_i)=0$，若局中人 i 捐款 c_i，即 $s_i(c_i)=1$，于是 $\max(s_1, s_2)=1, u_i=1-c_i$，又若 i 不捐款，而 j 捐款，$s_j(c_j)=1$，因此仍有 $\max(s_1, s_2)=1$，但 $s_i(c_i)=0$，于是 $u_i=1$。

这是一个Bayes博弈，它除了拥有局中人、纯策略空间与盈利函数这三个元素之外，还拥有类型及信念 $P(\cdot)$ 这两个要素。因此该博弈的Bayes均衡是一对策略 $(s_1^*(c_1), s_2^*(c_2))$，它使得对每一个局中人 i 及每一个可能的类型 c_i，$s_i^*(c_i)$ 使期望盈利 $E_{c_j}u_i(s_i, s_j^*(c_j), c_i)$ 极大化。这里记号 E_{c_j} 表示关于随机变量 c_j 求数学期望。不妨令 $z_j=\text{Prob}(s_j^*(c_j)=1)$（这儿概率不用 P 而用Prob，是为了不

与 $P(\cdot)$混淆)为局中人 j 捐款的(均衡)概率,现在我们来求期望盈利。

$$E_{c_j}u_i(s_i,s_j^*(c_j),c_i)=E_{c_j}\{\max(s_i(c_i),s_j^*(c_j))-c_is_i\}$$
$$=z_j\max(s_i(c_i),1)+(1-z_j)\max(s_i(c_i),0)-c_is_i$$
$$=z_j+(1-z_j)\max(s_i(c_i),0)-c_is_i \qquad (10.17)$$

当 $s_i(c_i)=1$(局中人 i 捐款)时,上式等于 $1-c_i$,而当 $s_i(c_i)=0$(局中人 i 不捐款)时,上式等于 z_j。因此只有 $1-c_i>z_j$ 或 $c_i<1-z_j$ 时局中人 i 才会采取捐款策略行动。于是,从极大化局中人 i 的期望盈利的角度,应有

$$\begin{cases}s_i^*(c_i)=1 & 若\ c_i<1-z_j\\ s_i^*(c_i)=0 & 若\ c_i>1-z_j\end{cases} \qquad (10.18)$$

那么,$c_i=1-z_j$ 时怎样呢?此时显然局中人 i 在捐款与不捐款问题上表现出无所谓的态度,因为这将给他带来相同的收益。但是由于我们已经假设类型 c_i 服从连续分布 $P(\cdot)$,因此 c_i 取特殊值 $1-z_j$ 的概率为 0,就是说,我们几乎不必考虑这种情况。虽然迄今为止我们尚不知道 z_j 到底等于多少,但(10.18)式告诉我们愿捐款的局中人 i 的类型有一个上界 c_i^*,仅当 $c_i\in[\underline{c},c_i^*]$(即局中人 i 的花费充分低)时他捐款 c_i。由于博弈对于双方是对称的,因此同样地可以知道,仅当 $c_j\in[\underline{c},c_j^*]$时局中人 j 捐款 c_j。

因为 $c_j\in[\underline{c},c_j^*]$时方有 $s_j^*(c_j)=1$ 故有

$$z_j=\text{Prob}(s_j^*(c_j)=1)=\text{Prob}(\underline{c}\leqslant c_j\leqslant c_j^*)=P(c_j^*)$$

由上述分析,均衡上限 c_i^* 必须满足 $c_i^*=1-z_j=1-P(c_j^*)$,因此,事实上 c_i^* 与 c_j^* 均满足方程

$$c^*=1-P(1-P(c^*)) \qquad (10.19)$$

倘若(10.19)式存在唯一解 c^*,那么必有 $c_i^*=c_j^*=c^*=1-P(c^*)$。例如,设 P 为[0,2]上的均匀分布,即对任何 $c\in[0,2]$,应有 $P(c)=c/2$,那么 $c^*=1-c^*/2$,推得 $c^*=2/3$。人们很可能存在

这样的想法，只要捐款小于1，那么他一定有正收益，于是捐款是划得来的。其实不然，Bayes 均衡告诉我们，即使捐款额位于(2/3，1]，它显然小于从公共财产得到的部分，此时局中人仍不捐款。

例 10.2 消耗战(War of Attrition)

考虑消耗战的不完全信息形式。两个局中人同时行动，局中人 i 从 $[0,+\infty)$ 中选取 $s_i(i=1,2)$。相应盈利函数为

$$u_i=\begin{cases}-s_i & \text{如果 } s_j\geqslant s_i\\ \theta_i-s_j & \text{如果 } s_j<s_i\end{cases} \tag{10.20}$$

其中 θ_i 是由于 s_i 高于 s_j 而使局中人 i 作为赢家所获得的奖励，它取值于 $[0,+\infty)$ 之中，作为局中人 i 的类型(因此是局中人 i 的私人信息)，具有 $[0,+\infty)$ 上累积分布 P 与密度 p。

现在寻求博弈的纯策略 Bayes 均衡 $(s_1(\cdot),s_2(\cdot))$，对于每一个类型 θ_i，$s_i(\theta_i)$ 必须满足使下式达到极大化：

$$-s_i\text{Prob}(s_j(\theta_j)\geqslant s_i)+\int_{\{\theta_j|s_j(\theta_j)<s_i\}}(\theta_i-s_j(\theta_j))p_j(\theta_j)d\theta_j \tag{10.21}$$

(10.21)式实际上是局中人 i 的期望盈利。$s_i(\theta_i)$ 作为类型 θ_i 的函数，如果是均衡策略，那么对于 θ_i'，必定要求与相应的 $s_i(\theta_i')=s_i'$ 一起代入(10.21)式才能使该式达到极大。换句话说，倘若以 $(\theta',s_i''=s_i(\theta_i''))$，其中 $(\theta_i'\neq\theta_i'')$，代入(10.21)式并不能使该式达到极大。若用不等式表述这段话的意思应有

$$\begin{aligned}&-s_i'\text{Prob}(s_j(\theta_j)\geqslant s_i')+\int_{\{\theta_j|s_j(\theta_j)<s_i'\}}(\theta_i'-s_j(\theta_j))p_j(\theta_j)\mathrm{d}\theta_j\\ \geqslant&-s_i''\text{Prob}(s_j(\theta_j)\geqslant s_i'')+\int_{\{\theta_j|s_j(\theta_j)<s_i''\}}(\theta_i'-s_j(\theta_j))p_j(\theta_j)\mathrm{d}\theta_j\end{aligned} \tag{10.22}$$

(10.22)式可写成

$$\theta_i'\text{Prob}(s_j(\theta_j)<s_i')-s_i'\text{Prob}(s_j(\theta_j)$$

$$\geqslant s_i')-\int_{\{\theta_j|s_j(\theta_j)<s_i'\}}s_j(\theta_j)p_j(\theta_j)\mathrm{d}\theta_j$$
$$\geqslant\theta_i'\mathrm{Prob}(s_j(\theta_j)<s_i'')-s_i''\mathrm{Prob}(s_j(\theta_j)$$
$$\geqslant s_i'')-\int_{\{\theta_j|s_j(\theta_j)<s_i''\}}s_j(\theta_j)p_j(\theta_j)\mathrm{d}\theta_j \qquad (10.23)$$

同样地，在上式中可以将 θ_i' 与 θ_i''，s_i' 与 s_i''互换，得

$$\theta_i''\mathrm{Prob}(s_j(\theta_j)<s_i'')-s_i''\mathrm{Prob}(s_j(\theta_j)$$
$$\geqslant s_i'')-\int_{\{\theta_j|s_j(\theta_j)<s_i''\}}s_j(\theta_j)p_j(\theta_j)\mathrm{d}\theta_j$$
$$\geqslant\theta_i''\mathrm{Prob}(s_j(\theta_j)<s_i')-s_i'\mathrm{Prob}(s_j(\theta_j)$$
$$\geqslant s_i')-\int_{\{\theta_j|s_j(\theta_j)<s_i'\}}s_j(\theta_j)p_j(\theta_j)\mathrm{d}\theta_j \qquad (10.24)$$

"(10.23)式的左端减去(10.24)式的右端"显然大于"(10.23)式的右端减去(10.24)式的左端"，于是我们有

$$(\theta_i''-\theta_i')\{\mathrm{Prob}(s_j(\theta_j)\geqslant s_i')-\mathrm{Prob}(s_j(\theta_j)\geqslant s_i'')\}\geqslant 0 \qquad (10.25)$$

由(10.25)式，如果 $\theta_i''\geqslant\theta_i'$，那么必有

$$\mathrm{Prob}(s_j(\theta_j)\geqslant s_i')\geqslant\mathrm{Prob}(s_j(\theta_j)\geqslant s_i'')$$

这表示 $s_i''\geqslant s_i'$。就是说，如果 $s_i(\theta_i)$是不完全信息形式消耗战博弈的 Bayes 均衡策略，那么 $s_i(\theta_i)$必定是类型 θ_i 的非减函数。事实上如果运用严格的数学证明，我们可以知道 $s_i(\theta_i)$是 θ_i 的严增函数，这里将略去证明而请读者记住结论。$s_i(\theta_i)$既然是严增函数，那么它一定存在逆函数 $\theta_i=\Phi_i(s_i)$。于是期望盈利(10.21)式可以写作

$$-s_i[1-\mathrm{Prob}(s_j(\theta_j)<s_i)]+\int_0^{s_i}(\theta_i-s_j)p_j(\Phi_j(s_j))\Phi_j'(s_j)\mathrm{d}s_j$$
$$=-s_i[1-P_j(\Phi_j(s_i)]+\int_0^{s_i}(\theta_i-S_j)p_j(\Phi_j(s_j))\Phi_j'(s_j)\mathrm{d}s_j \qquad (10.26)$$

注意，(10.26)式的获得只要通过在(10.21)式的积分中作变量代换 $\theta_j=\Phi_j(s_j)$，此时 $\mathrm{d}\theta_j=\Phi_j'(s_j)\mathrm{d}s_j$，$p_j(\theta_j)=p_j(\Phi_j(s_j))$。为使

(10.26)式达到极大，对 s_i 求导并使之等于 0 得到一阶条件

$$-[1-P_j(\Phi_j(s_i))]+s_i p_j(\Phi_j(s_i))\cdot\Phi_j'(s_i)$$
$$+(\theta_i-s_i)p_j(\Phi_j(s_i))\cdot\Phi_j'(s_i)=0$$

对上式进行整理并以 $\theta_i=\Phi_i(s_i)$ 代入上式，则有

$$\Phi_i(s_i)p_j(\Phi_j(s_i))\cdot\Phi_j'(s_i)=1-P_j(\Phi_j(s_i)) \qquad (10.27)$$

利用(10.27)式求得 Bayes 均衡策略 $s_i(\theta_i)$。不妨考虑一些特殊的情况，令 $P_1=P_2=P$，显然我们寻求的是对称 Bayes 均衡。在(10.27)式中，令 $\Phi(s)=\theta$，并且利用逆函数理论，我们有 $\Phi'(s)\cdot s'(\theta)=1$，即 $\Phi'(s)=1/s'$，于是(10.27)式变成

$$\theta\cdot p(\theta)\cdot\frac{1}{s'(\theta)}=1-P(\theta)$$

或者

$$s'(\theta)=\theta\cdot p(\theta)/[1-P(\theta)] \qquad (10.28)$$

注意到 $s(0)=0$ 这一事实(没有奖励也就不愿为之奋斗)，我们解得

$$s(\theta)=\int_0^\theta\frac{xp(x)}{1-P(x)}\mathrm{d}x \qquad (10.29)$$

例 10.3 拍卖

考虑一次价格密封标价拍卖(first-price, sealed-bid auction)。两个标价者，记作 $i=1,2$。对于拍卖的货物，每个标价者 i 有自己的估价 v_i，如果他以价格 p 获得货物，那么他将获益 v_i-p。现假设 v_1 与 v_2 为来自均匀分布 $R[0,1]$ 上的独立随机变量。对于标价有一个合理的约束：标价不可为负数。标价者以密封形式投标。谁开价高，他就获得待售货物并为此付出自己的标价；另一个标价者得不到货也不必为此掏腰包(相比货品的价值，我们忽略了那些相对来说极微小的拍卖费用)。倘若两个人标价恰好相等，那么只能通过抛一枚均匀硬币来确定谁是赢者。以上所述视作博弈的共同知识。

现在我们对一次价格密封拍卖建立博弈模型。

(1)行动空间——局中人 i 的行动是递送一个(非负)标价 b_i，$b_i \in A_i = [0, +\infty)$。

(2)类型空间——局中人 i 的类型是他对货物的估价 v_i，由假设，类型空间 $T_i = [0, 1]$。

(3)信念——估价是独立的，因此局中人 i 相信 v_j 均匀地分布在$[0,1]$上。

(4)盈利函数

$$u_i(b_1, b_2, v_1, v_2) = \begin{cases} v_i - b_i & \text{如果 } b_i > b_j \\ (v_i - b_i)/2 & \text{如果 } b_i = b_j \\ 0 & \text{如果 } b_i < b_j \end{cases} \tag{10.30}$$

(10.30)式中，$b_i = b_j$ 时的盈利函数表示的是期望盈利。

在这个 Bayes 博弈中，局中人 i 的策略应当是类型 v_i 的函数，记为 $b_i(v_i)$。根据定义，Bayes 均衡要求标价者 1 的策略 $b_1(v_1)$是关于标价者 2 的策略 $b_2(v_2)$的最佳反应，标价者 2 的 $b_2(v_2)$也是关于 $b_1(v_1)$最佳反应。若$(b_1(v_1), b_2(v_2))$是 Bayes 均衡，那么对于每一个 $v_i \in A_i = [0, 1]$，$(i = 1, 2)$，$b_i(v_i)$必须使下式极大化：

$$(v_i - b_i)\text{Prob}\{b_i > b_j(v_j)\} + \frac{1}{2}(v_i - b_i)\text{Prob}\{b_i = b_j(v_j)\} \tag{10.31}$$

不从数学的角度而完全地凭直觉，一般地人们都会接受这样一种看法：对货物的估价 v 越高，常常标价 b 也会随之增加，简而言之，$b_i(v_i)$应当是 v_i 的非减函数。假设 $b_i(v_i)$是较光滑的函数，我们不妨以其 Taylor 展开的第一阶作为近似。换句话说，我们试图寻求线性 Bayes 均衡：$b_1(v_1) = a_1 + c_1 v_1$ 和 $b_2(v_2) = a_2 + c_2 v_2$，从而对我们的问题简化地进行阐述。当然寻求线性均衡不等于限制局中人的策略空间为线性空间。

首先注意到，由于 $b_j(v_j) = a_j + c_j v_j$ 以及 v_j 服从$[0,1]$上的均

匀分布，$b_j(v_j)$服从$[a_j, a_j+c_j]$上的均匀分布，必有 $\text{Prob}\{b_i=b_j(v_j)\}=0$，故对于每一个类型 v_i，局中人 i 的最佳反应 $b_i(v_i)$一定使下式极大化：

$$(v_i-b_i)\text{Prob}\{b_i>b_j(v_j)\} \tag{10.32}$$

理性告诉我们，b_i 应满足 $a_j\leqslant b_i\leqslant a_j+c_j$，就是说，局中人 i 的标价应在标价者 j 的最小可能标价之上，否则其标价将显得毫无意义，同时标价者 i 的标价也应该不超过标价者 j 的最大可能标价，要不，他是一个极愚蠢的人。这样，

$$\text{Prob}\{b_i>a_j+c_jv_j\}=\text{Prob}\{v_j<\frac{b_i-a_j}{c_j}\}=\frac{b_i-a_j}{c_j} \tag{10.33}$$

期望盈利为

$$\frac{(v_i-b_i)(b_i-a_j)}{c_j}=\frac{[-b_i^2+(a_j+v_i)b_i+v_ia_j]}{c_j}$$

如果利用一阶条件，易知局中人 i 的最佳反应为 $b_i=(v_i+a_j)/2$，但由于如刚才所述，b_i 必须大于等于 a_j，因此倘若 $v_i<a_j$ 的话，必须至少取 $b_i=a_j$。明确地用公式表示如下：

$$b_i(v_i)=\begin{cases}(v_i+a_j)/2 & \text{若 } v_i\geqslant a_j\\ a_j & \text{若 } v_i<a_j\end{cases} \tag{10.34}$$

$b_j(v_j)=a_j+c_jv_j$ 中的系数是假设的，倘若 $0<a_j<1$，那么 $v_i<a_j$ 的确有可能，但此时由(10.34)式可知，$b_i(v_i)$ 就不是 v_i 的线性函数．因为当 $v_i<a_j$ 时 $b_i(v_i)$是一条平坦的直线，而当 $v_i\geqslant a_j$ 时，$b_i(v_i)$则开始向上倾斜，整体成为一条折线，这与原先假设 $b_i(v_i)$的线性函数存在着矛盾。或者说，如果 $a_j\in(0,1)$，那么就不可能存在线性 Bayes 均衡 $b_i(v_i)$，由于我们目前寻求的是线性均衡，因此我们不考虑 $0<a_j<1$，而代之于关心 $a_j\geqslant 1$ 与 $a_j\leqslant 0$ 的情况。但是当$a_j\geqslant 1$时，$b_j(v_j)=a_j+c_jv_j$ 已经假设 $b_j(v_j)$ 随 v_j 增加而增加，故 $c_j\geqslant 0$。于是 $a_j\geqslant 1\geqslant v_j$ 蕴含着 $b_j(v_j)\geqslant v_j$，这对于标价者 j 来说，看来不是一个最佳策略．于是，如果 $b_i(v_i)$ 是线性均衡的话，必有 $a_j\leqslant 0$。此时

成立 $b_i(v_i)=(v_i+a_j)/2=a_i+c_iv_i$，注意到 v_i 是个变量，立得 $a_i=a_j/2, c_i=1/2$。

在上面的讨论中，如果将 i 与 j 互相置换，结论当然成立。因此也有 $a_j=a_i/2, c_i=1/2$。于是，必有

$$a_i=a_j=0, c_i=c_j=1/2, b_i(v_i)=v_i/2, b_j(v_j)=v_j/2$$

我们在信念为均匀分布的条件下求得了一次价格密封拍卖的线性 Bayes 均衡，每个标价者递交的标价是他自己关于货物估价的一半。每一个赢者总是赢得他自己的标价，而标价越高当然越有可能赢，但这依赖它的类型——关于货物的估价。很自然会提出如下问题：(1)除了线性均衡之外，是否还有其他形式的 Bayes 均衡？(2)倘若估价的分布发生改变而不再是均匀分布，那么均衡标价如何随之变化？

假如我们试图去试试任何其他形式的策略函数，也许没多大意思，我们的确也无能对一切函数形式一一检验，因为这实在超出了我们的能力。而若估价分布换成其他累积分布函数，则将不存在线性均衡，这是因为此时(10.33)式的第二个等式后面不再是 b_i 的一次函数，因而期望盈利就不是 b_i 的二次型式的缘故。对于一般情况，我们仅限于讨论局中人策略为恒同的对称 Bayes 均衡。也就是说，存在一个函数 $b(\cdot)$ 使得局中人 i 的策略函数为 $b(v_i)(i=1,2)$。

现在我们假定 $b(\cdot)$是严格增加且可微的函数，于是局中人 i 的最佳标价当使如下期望盈利极大化：

$$(v_i-b_i)\mathrm{Prob}\{b_i>b(v_j)\} \tag{10.35}$$

这里简记 $b_i=b(v_i)$，由于 $b(\cdot)$的严格单调性，因此 $v_i=b^{-1}(b_i)$，b^{-1}表示函数 $b(\cdot)$的逆函数。仍假定 $v_i(i=1,2)$服从$[0,1]$上的均匀分布，从而

$$\mathrm{Prob}\{b_i>b(v_j)\}=\mathrm{Prob}\{b^{-1}(b_i)>v_j\}=b^{-1}(b_i) \tag{10.36}$$

将(10.36)式代人(10.35)式得期望盈利函数

$$(v_i-b_i)b^{-1}(b_i) \quad (10.37)$$

由于 $b(\cdot)$ 严增且可微，故逆函数 $b^{-1}(\cdot)$ 也具有同样良好的性质，在(10.37)式中对 b_i 求导并使之等于 0，我们得到局中人 i 最优标价的一阶条件：

$$-b^{-1}(b_i)+(v_i-b_i)\frac{\mathrm{d}b^{-1}(b_i)}{\mathrm{d}b_i}=0 \quad (10.38)$$

注意到我们在这里探索的是对称 Bayes 均衡，因此 $b_i(v_i)=b(v_i)$，即具有与 b_j 同样的函数形式，于是将 $b_i=b(v_i)$ 代入一阶条件(10.38)式，并注意 $b^{-1}(b_i)=b^{-1}(b(v_i))=v_i$ 这一事实，(10.38)式就成为

$$-v_i+(v_i-b(v_i))\cdot\frac{\mathrm{d}b^{-1}(b_i)}{\mathrm{d}b_i}=-v_i+(v_i-b(v_i))\cdot\frac{1}{b'(v_i)}=0 \quad (10.39)$$

或者

$$b'(v_i)\,v_i+b(v_i)=v_i \quad (10.40)$$

(10.40)式等价于

$$\frac{\mathrm{d}\{b(v_i)v_i\}}{\mathrm{d}v_i}=v_i \quad (10.41)$$

对(10.41)式两端积分得

$$b(v_i)v_i=\frac{1}{2}v_i^2+k \quad (10.42)$$

其中 k 是待定常数，我们需要一些初始条件或者限制条件以确定 k 的取值。毋庸置疑，没有一个标价者会干出递交比自己的估价更高标价的蠢事，即 $b(v_i)\leqslant v_i$，特别地取 $v_i=0$，并回想标价必须为非负，故

$$0\leqslant b(0)\leqslant 0 \qquad 即 \qquad b(0)=0 \quad (10.43)$$

由(10.43)式不难获得 $k=0$。因此，$b(v_i)=v_i/2$。这是一个有趣的结果，在信念服从均匀分布的条件下，我们放宽策略函数的线性假设，得到的 Bayes 均衡策略与线性假设时的 Bayes 均衡完全一致。

例 10.4 双边拍卖(Double Auction)

上面我们讨论了两个标价者参与拍卖竞争的博弈模型。现在考虑买卖双边的拍卖。双方对货物都有自己的估价作为私人信息。卖者提出要价 p_1，买者同时提出一个买价 p_2。如果 $p_2 \geqslant p_1$，那么交易发生并确定交易价为 $p=(p_1+p_2)/2$；如果 $p_1>p_2$，那么不发生交易。

卖者(局中人 1)对货物有一个估价 c，买者(局中人)的估价为 v，c 与 v 均属于区间$[0,1]$。如果买者以价格 p 得到货物，他的盈利为 $u_2=v-p$；如果不成交，则 $u_2=0$。倘若卖者以价格 p 出售货物，他的盈利为 $u_1=p-c$，如果不成交，则 $u_1=0$。这里盈利函数事实上测定了各方盈利的变化，因此在不成交的时候，盈利没有发生变化，故指定为 0。

设买者的策略是函数 $p_2(v)$，它确定了对于买者每一个可能的估价 v，他所愿意开出的买价。卖者的策略是函数 $p_1(c)$，它确定了对于卖者的每一个估价 c，他需求卖出的价格。如果要使 $\{p_1(c), p_2(v)\}$ 是 Bayes 均衡，则对于$[0,1]$中每一个 c，$p_1(c)$ 应使下式

$$\left\{\frac{p_1+E[p_2(v)\mid p_2(v)\geqslant p_1]}{2}-c\right\}\cdot \text{Prob}\{p_2(v)\geqslant p_1\} \tag{10.44}$$

达到极大化。其中 $E[p_2(v)\mid p_2(v)\geqslant p_1]$表示在买者的开价大于卖者需求 p_1 的条件下，买者将要开出的期望价格。同时，对于每一个 $v\in[0,1]$，$p_2(v)$应使

$$\left\{v-\frac{p_2+E[p_1(c)\mid p_2\geqslant p_1(c)]}{2}\right\}\cdot \text{Prob}\{p_2\geqslant p_1(c)\} \tag{10.45}$$

极大化。其中 $E[p_1(c)\mid p_2\geqslant p_1(c)]$表示在卖者的需求小于买者开价 p_2 的条件下，卖者需求的期望价格。(10.44)式与(10.45)式这

两个期望盈利是怎样来的呢？不妨以(10.44)式为例进行计算，由模型可知，卖者的盈利函数为

$$u_1=\begin{cases}\dfrac{p_1+p_2}{2} & 当\ p_1\leqslant p_2\\ 0 & 当\ p_1>p_2\end{cases} \tag{10.46}$$

对于每一个 $c\in[0,1]$，以及相应的 $p_1(c)=p_1$，由于卖者不知道 v，因此成交时的平均(或期望)盈利应为

$$\int_{p_1}^{1}\left(\frac{p_1+p_2}{2}-c\right)\mathrm{d}F_2(p_2) \tag{10.47}$$

其中

$$F_2(p_2)=\mathrm{Prob}(p_2(v)<p_2) \tag{10.48}$$

注意到

$$\begin{aligned}\int_{p_1}^{1}\mathrm{d}F_2(p_2)&=F_2(1)-F_2(p_1)\\ &=\mathrm{Prob}(p_2(v)\leqslant 1)-\mathrm{Prob}(p_2(v)<p_1)\\ &=\mathrm{Prob}(p_2(v)\geqslant p_1)\end{aligned} \tag{10.49}$$

且

$$\int_{p_1}^{1}p_2\mathrm{d}F_2(p_2)=E(p_2(v)\mid p_2(v)\geqslant p_1]\mathrm{Prob}(p_2(v)\geqslant p_1) \tag{10.50}$$

将(10.49)式、(10.50)式代入(10.47)式，就可以得到(10.44)式。现在我们回头来看(10.48)式，$F_2(p_2)$ 表示买者具有某些估价 v 的概率，这些 v 导致他开出低于 p_2 的价格。同样地我们可以用 $F_1(p_1)$表示卖者拥有导致他开出需求价低于 p_1 的成本 c 的概率。利用 $F_1(p_1)$，我们可以类似地获悉(10.45)式的由来。

如果 $p_1(c)$ 是 Bayes 均衡策略，那么对于 c'及 $p_1'=p_1(c')$，我们应有

$$\int_{p_1}^{1}\left(\frac{p_1+p_2}{2}-c\right)\mathrm{d}F_2(p_2)\geqslant\int_{p_1'}^{1}\left(\frac{p_1'+p_2}{2}-c\right)\mathrm{d}F_2(p_2)\tag{10.51}$$

$$\int_{p_1'}^{1}\left(\frac{p_1'+p_2}{2}-c'\right)\mathrm{d}F_2(p_2)\geqslant\int_{p_1}^{1}\left(\frac{p_1+p_2}{2}-c'\right)\mathrm{d}F_2(p_2)\tag{10.52}$$

“(10.51)式的左端减去(10.52)式右端”应该大于等于“(10.51)式的右端减去(10.52)式的左端”，并经整理得

$$\int_{p_1}^{1}(c'-c)\mathrm{d}F_2(p_2)\geqslant\int_{p_1'}^{1}(c'-c)\mathrm{d}F_2(p_2)$$

或者

$$(c'-c)\int_{p_1}^{p_1'}\mathrm{d}F_2(p_2)\geqslant 0\tag{10.53}$$

(10.53)式蕴含了，如果 $c'>c$，必有 $p_1'\geqslant p_1$，即均衡策略 $p_1(c)$ 是 c 的非减函数，同理可证，均衡策略 $p_2(v)$ 是 v 的非降函数。这与实际情况非常吻合，估价越高自然开价也就越高。

不完全信息的双边拍卖博弈可能有着许多 Bayes 均衡，我们仅考虑特殊的均衡，例如简单地我们仍考虑线性 Bayes 均衡；同样，这种考虑并不等于限制局中人的策略空间为线性空间，相反地，我们允许局中人任意地选择各种函数形式的策略，只不过在这里将探索是否存在一个均衡是线性形式的。

在探索之前，我们再作一个假定，v 与 c 为来自[0,1]上均匀分布的独立变量。现在假设卖者的策略函数为 $p_1(c)=a_1+e_1c$，于是 p_1 在$[a_1,a_1+e_1]$上均匀地分布。观察买者的期望盈利

$$\left\{v-\frac{p_2+E[p_1(c)\mid p_2\geqslant p_1(c)]}{2}\right\}\mathrm{Prob}\{p_2\geqslant p_1(c)\}$$
$$=\int_0^{p_2}\left(v-\frac{p_1+p_2}{2}\right)\mathrm{d}F_1(p_1)$$
$$=\left(v-\frac{p_2}{2}\right)\int_0^{p_2}\mathrm{d}F_1(p_1)-\frac{1}{2}\int_0^{p_2}p_1\mathrm{d}F_1(p_1)\tag{10.54}$$

由于

$F_1(p_1)=\text{Prob}(p_1(c)\leqslant p_1)=\frac{p_1-a_1}{e_1}\quad(p_1\geqslant a_1)$，$d(F_1(p_1))=\frac{1}{e_1}\mathrm{d}p_1$，故

$$(10.54)\text{式}=(v-\frac{p_2}{2})F_1(p_2)-\frac{1}{2}\int_{a_1}^{p_2}\frac{p_1}{e_1}\mathrm{d}p_1$$
$$=[v-\frac{1}{2}(p_2+\frac{p_2+a_1}{2})]\frac{p_2-a_1}{2e_1}\quad(10.55)$$

(10.54)式中的第一个等号的证明类似于(10.47)式的获得，我们不再赘述。使(10.55)式极大化的一阶条件为

$$p_2=\frac{2}{3}v+\frac{1}{3}a_1\quad(10.56)$$

(10.56)式指出，如果信念为均匀分布，且如果卖者取线性策略的话，那么买者的最佳反应也是线性函数。类似地，假使买者取线性策略 $p_2(v)=a_2+e_2v$，那么 p_2 在$[a_2,a_2+e_2]$上均匀地分布，利用(10.47)式卖者的期望盈利为

$$\int_{p_1}^{1}(\frac{p_1+p_2}{2}-c)\mathrm{d}F_2(p_2)$$
$$=\{\frac{1}{2}[p_1+\frac{p_1+a_2+e_2}{2}]-c\}\frac{a_2+e_2-p_1}{e_2}$$
$$(10.57)$$

使(10.57)式极大化的一阶条件应为

$$p_1=\frac{2}{3}c+\frac{1}{3}(a_2+e_2)\quad(10.58)$$

显然，此时卖者的最佳反应也是线性函数。

如果买卖双方的线性策略互为最佳反应，对比系数立即可得

$$p_2(v)=\frac{2}{3}v+\frac{1}{12}\quad(10.59)$$

$$p_1(c)=\frac{2}{3}c+\frac{1}{4}\quad(10.60)$$

买者的最大开价 $p_2(1)=\frac{2}{3}+\frac{1}{12}=\frac{9}{12}=\frac{3}{4}$。如果 $c>\frac{3}{4}$的话，那么卖者的需求价小于他的成本 c，因为此时$\frac{1}{4}<\frac{c}{3}$，故有 $p_1(c)<\frac{2}{3}c+\frac{1}{3}c=c$。很清楚，当成本比较大时，纵然卖者开价小于他的成本价，但是仍大于买者的最大开价，此时不会发生交易，这件事实表明，在 Bayes 均衡中，卖者不会以低于成本价的价格出售自己的货物。同样地，令$\frac{2}{3}v+\frac{1}{12}>v$ 推得 $v<\frac{1}{4}$，即当 $v<\frac{1}{4}$时，买者对货物的估价相当低，按均衡策略，此时卖者的开价超过了自己对货物的估价，当 $0<v<\frac{1}{4}$，开价小于$\frac{1}{4}$，这个开价小于卖者最小的需求价 $p_1(0)=\frac{1}{4}$，因此也不会发生交易，这说明，在 Bayes 均衡中，买者也不可能以超出自己估价的价格购买货物。

模型的假设为，当且仅当 $p_2(v)\geqslant p_1(c)$时才会使双边拍卖有成交的可能。也就是当且仅当 $v\geqslant c+\frac{1}{4}$时交易以线性均衡形式发生。这与我们平时所想象的，当 $v\geqslant c$ 时会发生交易，交易的范围小了一点。然而人们不难发现，Bayes 线性均衡的条件使卖者“有利可图”的可能性增大。

§10.5 混合策略的再解释

在完全信息静态博弈中，我们已经引进了混合策略均衡这个重要的概念。但是许多研究人员并不喜欢混合策略，简单且重要的理由在于：在现实世界中，不可能通过抛一枚硬币来作出重要的甚至性命攸关的决策。混合策略的另一种解释是它表示了局中人将要采取的行动的不确定性。1973 年，Harsanyi 指出，完全信息博弈的混合策略均衡通常可以解释为稍微受到扰动的不完全信息博弈

的纯策略均衡的极限。这是因为，Harsanyi 认为不确定性归因于有关对手的盈利的少量不确定性。所谓不完全信息博弈受到稍许扰动，就是指局中人的盈利有少量的幅度变化。我们将以若干例子详细地阐述混合策略的这类再解释。

例 10.5 "抓钱"(Grab tbe Dollar)

考虑"抓钱"博弈，每个局中人有两个可能行动：投资("抓")或不投资。在完全信息静态博弈中两个人盈利矩阵如图 10.5。

	抓	不抓
抓	−1，−1	1，0
不抓	0，1	0，0

图 10.5

该博弈的唯一对称均衡是每个局中人以$\frac{1}{2}$概率抓钱(如换成其他背景，每个公司以$\frac{1}{2}$概率投资)。这是因为每个公司若不投资则显然获益为 0，而如果它投资的话，其期望盈利为$\frac{1}{2}(1)+\frac{1}{2}(-1)=0$，因此公司对是否投资无所谓，于是混合策略$(\frac{1}{2},\frac{1}{2})$是 Nash 均衡。

现在我们稍微扰动一下盈利矩阵，即对部分盈利赋予一个"小小的"随机干扰，以使抓钱博弈具有不完全信息的类型：每家公司成为赢家时其盈利为$(1+\theta_i)(i=1,2)$、其中 θ_i 在 $[-\epsilon,\epsilon]$ 上均匀地分布，ϵ 是一个相当小的正数。θ_i 是公司 i 的私人信息，即类型，只有公司 i 知道而不为其他公司知晓。用正则型式表示为图 10.6。

这显然是个不完全信息的 Bayes 博弈，不管公司属于何种类型，它们仍然各有两个策略：投资或不投资。信念密度为 $p_1(\theta_2)=p_2(\theta_1)=\frac{1}{2\epsilon}$。由于决策为 0—1 型式，因此当 θ_1 超过某临界值 x 时，

公司 2

		投资	不投资
公司 1	投资	$-1,-1$	$1+\theta_1,0$
	不投资	$0,1+\theta_2$	$0,0$

图 10.6

公司 1 投资，否则就不投资，同样，如果 θ_2超过某临界值 y，公司 2 投资，否则就不投资。$x,y\in[-\varepsilon,\varepsilon]$ 。考虑公司 1，$P(\theta_1<x)=\frac{x+\varepsilon}{2\varepsilon}$，即公司 1 以概率$\frac{x+\varepsilon}{2\varepsilon}$不投资，而以概率$\frac{\varepsilon-x}{2\varepsilon}$投资。同样，公司 2 以概率$\frac{\varepsilon-y}{2\varepsilon}$投资，而以概率$\frac{y+\varepsilon}{2\varepsilon}$不投资。不难计算公司 1 投资时的期望盈利等于

$$(-1)\cdot\frac{\varepsilon-y}{2\varepsilon}+(1+\theta_1)\frac{y+\varepsilon}{2\varepsilon}=\frac{y}{\varepsilon}+\frac{\theta_1(y+\varepsilon)}{2\varepsilon} \quad (10.61)$$

而不投资的盈利仍为 0，即当且仅当

$$\frac{y}{\varepsilon}+\frac{\theta_1(y+\varepsilon)}{2\varepsilon}\geqslant 0$$

或

$$\theta_1\geqslant-\frac{2y}{y+\varepsilon}=x \quad (10.62)$$

时公司 1 投资，类似地，当且仅当

$$\theta_2\geqslant-\frac{2x}{x+\varepsilon}=y \quad (10.63)$$

时公司 2 投资，同时解(10.62)式与(10.63)式，x 与 y 的合理解必为 $x=y=0$。于是我们得到对称的纯策略 Bayes 均衡解为：公司 i 当 $\theta_i\geqslant 0$ 时投资，当 $\theta_i<0$ 时不投资$(i=1,2)$。有趣的是，不管正 ε 大约取多少，每家公司均以$\frac{1}{2}$概率投资，$\frac{1}{2}$概率不投资。而当 $\varepsilon\to 0$ 时，图 10.6 趋于图 10.5，可以这样认为，当 ε 趋于零时，不完全信息静态博弈趋于完全信息静态博弈，而完全信息静态博弈的混合

策略 Nash 均衡$((\frac{1}{2},\frac{1}{2}),(\frac{1}{2},\frac{1}{2}))$可以视作这一系列 Bayes 博弈的 Bayes 均衡的极限。

用抓钱博弈来再解释混合策略均衡，使读者会似有所悟，但也许会有读者提出，这个例子太特殊了，因为每个受到扰动之后形成的 Bayes 博弈的 Bayes 均衡恰好就是完全信息静态博弈的混合策略均衡解。我们将用其他一些例子进一步阐述"再解释"问题。

例 10.6 性别战

这里不再详述性别战的具体内容，只再一次展示性别战的盈利矩阵(见图 10.7)。

		丈夫	
		戏	足球
妻子	戏	2,1	0,0
	足球	0,0	1,2

图 10.7

该博弈有两个纯策略 Nash 均衡(戏，戏)与(足球，足球)，另外还有一个混合策略 Nash 均衡$((\frac{2}{3},\frac{1}{3}),(\frac{1}{3},\frac{2}{3}))$。

为了对图 10.7 盈利矩阵稍稍赋予扰动，不妨作一个假设，夫妻俩尽管已经互相了解有一段时期，但毕竟还不能完全肯定地把握对方的想法。特殊地，如果双方都取看戏，妻子的获益为 $2+t_1$，其中 t_1 为妻子的私人信息，尚未被丈夫知晓；如果两人都去看足球的话，丈夫的获益为 $2+t_2$，其中 t_2 为丈夫的私人信息。设类型 t_1 与 t_2 为来自$[0,x]$上均匀分布的两个独立变量，$[0,x]$上均匀分布的选择在这里不是重要的关键，我们的原意只是利用 t_1 与 t_2 稍稍干扰原来性别战中的盈利矩阵，故应认为 x 是个相当小的正数。新得到的静态 Bayes 博弈的盈利(或效用)矩阵如图 10.8 所示。

丈夫

		戏	足球
妻子	戏	$2+t_1$,1	0,0
	足球	0,0	$1,2+t_2$

图 10.8

在该 Bayes 博弈中，行动空间 $A_1=A_2=\{$戏，足球$\}$，类型空间 $T_1=T_2=[0,x]$，信念密度为 $p_1(t_2)=p_2(t_1)=1/x$ 对一切 t_1 与 t_2 成立。注意到，尽管性别战受到稍微干扰而成为不完全信息静态博弈，但是双方的策略仍为 0—1 决策的，因此，虽然类型空间是连续统区间$[0,x]$，显然 t_1 应有一个临界值 c_1，t_1 超过 c_1，妻子会选择行动——看戏，否则就取看足球，同样 t_2 也有一个临界值 c_2，当 t_2 超过 c_2 时，丈夫毫不犹豫地取看足球，否则取看戏。现在我们就来寻求 Bayes 博弈的纯策略均衡，在这个均衡中，妻子以 $(x-c_1)/x$ 概率看戏，而以 c_1/x 概率看足球；丈夫以 $(x-c_2)/x$ 概率看足球，而以 c_2/x 概率看戏。显然，摆在我们面前的任务是，对于给定的 x，求出 c_1 与 c_2 的确定值，以使上述策略构成纯策略 Bayes 均衡。

给定丈夫的策略，妻子看戏的期望盈利为

$$\frac{c_2}{x}(2+t_1)+\frac{x-c_2}{x}\cdot 0=\frac{c_2}{x}(2+t_1) \tag{10.64}$$

妻子看足球的期望盈利为

$$\frac{c_2}{x}\cdot 0+\frac{x-c_2}{x}\cdot 1=1-\frac{c_2}{x} \tag{10.65}$$

当且仅当

$$t_1\geqslant\frac{x}{c_2}-3=c_1 \tag{10.66}$$

时，妻子取看戏是关于丈夫策略的最佳反应。类似地，给定妻子的策略，丈夫看足球的期望盈利为

$$\left(1-\frac{c_1}{x}\right)\cdot 0+\frac{c_1}{x}(2+t_2)=\frac{c_1}{x}(2+t_2) \tag{10.67}$$

丈夫看戏的期望盈利为

$$\left(1-\frac{c_1}{x}\right)\cdot 1+\frac{c_1}{x}\cdot 0=1-\frac{c_1}{x} \tag{10.68}$$

当且仅当

$$t_2 \geqslant \frac{x}{c_1}-3=c_2 \tag{10.69}$$

时，丈夫取看足球是关于妻子策略的最佳反应。解(10.66)式与(10.69)式后，易得

$$\begin{cases} c_1=c_2 \\ c_2^2+3c_2-x=0 \end{cases} \tag{10.70}$$

解上述关于 c_2 的两次方程，并取其合理解，得

$$c_1=c_2=\frac{-3+\sqrt{9+4x}}{2} \tag{10.71}$$

因此，妻子看戏的均衡概率 $\frac{(x-c_1)}{x}$ 和丈夫看足球的概率 $\frac{(x-c_2)}{x}$ 都等于

$$1+\frac{3-\sqrt{9+4x}}{2x} \tag{10.72}$$

当 x 趋于 0 时，(10.72)式趋于 $\frac{2}{3}$。这表明，当不完全信息消失时 $(x\to 0)$、在不完全信息博弈中的纯策略 Bayes Nash 均衡中局中人的行为趋向于原来完全信息博弈中混合策略 Nash 均衡中这些局中人各自的行为。

上面两个例子是 0—1 决策型的，下面的例子为行动空间具有连续统势。

例 10.7 对称消耗战

假定下列盈利函数是共同知识：

$$u_i(s_i,s_j)=\begin{cases}-s_i & \text{如果 } s_j\geqslant s_i \\ \theta-s_j & \text{如果 } s_j<s_i\end{cases} \tag{10.73}$$

该博弈至少有两个非对称均衡。例如，以自然垄断解释，某公司永远占有，而另一个公司永远退出竞争就是一个非对称均衡，只要从盈利函数的取值容易证明这的确是个均衡。现在我们关心的是消耗战的对称均衡，这是一个混合策略均衡（这里 θ 为固定值）。既然行动空间是$[0,+\infty)$，因此所谓混合策略应当是$[0,+\infty)$上的累积分布函数；在 s 之前停止的概率分布函数为 $F(s)=1-\exp(-s/\theta)$，相应的密度为 $f(s)=\frac{1}{\theta}\exp(-s/\theta)$；于是在 s 之前不停止（概率为 $\exp(-s/\theta)$）条件下，在 s 与$(s+\mathrm{d}s)$之间局中人停止的概率为

$$\frac{\left[\frac{1}{\theta}\exp\left(-\frac{s}{\theta}\right)\right]\mathrm{d}s}{\exp(-s/\theta)}=\frac{\mathrm{d}s}{\theta} \tag{10.74}$$

为什么这个策略剖面是 Nash 均衡呢？考虑局中人多逗留 $\mathrm{d}s$ 所得到的期望获益为 $\theta\cdot\mathrm{d}s/\theta$，而多逗留 $\mathrm{d}s$ 的成本当然是 $\mathrm{d}s$，这两者显然相等，从而使局中人对在 $s+\mathrm{d}s$ 之前停止还是在 s 之前停止抱无所谓态度，这正是混合策略均衡的要点。或者可以这样地表述：在每一时刻，在两个局中人仍然争斗的条件下，每一个局中人从该时刻开始的期望估价等于 0。使得局中人在继续争斗还是退出争斗问题上感到无所谓，从而所提出的混合策略剖面是 Nash 均衡。

这个混合策略均衡能否通过一系列 Bayes 均衡提纯而得到？就是说，是否存在类型的一系列连续分布，它们弱收敛于质量在 θ 的一点（退化）分布，使得每一个类型取一个纯策略，且行动的均衡分布收敛于与完全信息博弈的混合策略均衡相关联的分布。

我们对博弈予以适当的扰动，此时 θ 就不是一个原有的固定值，而是一个在$[0,\infty)$上连续变化的变量，只不过其概率质量较集中于$(\theta-\epsilon,\theta+\epsilon)$（显然我们不会考虑奖励 $\theta=0$ 的情况，因为这

不会激励局中人投入消耗战，因此只要 ϵ 适当地小，$(\theta-\epsilon,\theta+\epsilon)$区间是现实的）。

注意到我们只考虑对称均衡，因此两个局中人的信念分布相同，考虑 $[0,\infty)$上的概率密度 $p^n(\cdot)$（随 n 的不同，概率密度因而也不同），相应的累积分布函数记作 $P^n(\cdot)$，它使得 $P^n(0)=0$，并且对所有 $\epsilon>0$，成立

$$\lim_{n\to\infty}[P^n(\theta+\epsilon)-P^n(\theta-\epsilon)]=1 \tag{10.75}$$

由于概率分布函数的特点，(10.75)蕴含着 $P^n(\theta-\epsilon)\to 0$，且 $P^n(\theta+\epsilon)\to 1$，也就是说，随着 n 的增加，分布函数序列 $P^n(\cdot)$越来越将概率质量集中于区间$(\theta-\epsilon,\theta+\epsilon)$。记 $s^n(\cdot)$为对应于信念 $P^n(\cdot)$的对称均衡策略，不妨记 Φ^n 为 s^n 的逆函数（由例 10.2，$s^n(\cdot)$是类型的严增函数，必存在逆函数）利用公式(10.27)应有

$$\Phi^n(s)P^n(\Phi^n(s))(\Phi^n(s))'=1-P^n(\Phi^n(s)) \tag{10.76}$$

将(10.76)式改写为

$$\frac{\mathrm{d}P^n(\Phi^n(s))}{1-P^n(\Phi^n(s))}=\frac{1}{\Phi^n(s)}$$

或者

$$\frac{\mathrm{d}(1-P^n(\Phi^n(s)))}{1-P^n(\Phi^n(s))}=-\frac{1}{\Phi^n(s)} \tag{10.77}$$

解(10.77)，易得

$$P^n(\Phi^n(s))=1-\exp\{-\int_0^s \mathrm{d}b/\Phi^n(b)\} \tag{10.78}$$

由于 $P^n(\theta-\epsilon)=P^n(\Phi^n(s^n(\theta-\epsilon)))\to 0$，对照(10.78)式，

$$P^n(\Phi^n(s^n(\theta-\epsilon)))=1-\exp\left(-\int_0^{s^n(\theta-\epsilon)}\mathrm{d}b/\Phi^n(b)\right)\to 0 \tag{10.79}$$

必有 $s^n(\theta-\epsilon)\to 0$ 对所有 $\epsilon<0$ 成立。又

$$P^n(\Phi^n(s^n(\theta+\varepsilon))) = 1 - \exp\left(-\int_0^{s^n(\theta+\varepsilon)} \mathrm{d}b/\Phi^n(b)\right) \to 1$$

(10.80)

推得

$$\exp\left(-\int_0^{s^n(\theta+\varepsilon)} \frac{\mathrm{d}b}{\Phi^n(b)}\right) \qquad (10.81)$$

当 $n\to\infty$ 时，对任意 $\varepsilon>0$，$s^n(\theta+\varepsilon)$ 只能趋于无限。因此，当 n 充分大时，对任意 $s>0$ 及 $\varepsilon\in(0,\theta)$，应有

$$s^n(\theta-\varepsilon)<s<s^n(\theta+\varepsilon)$$

于是(10.78)式可以写成

$$\begin{aligned} P^n(\Phi^n(s)) &= 1 - \exp\left(-\int_0^{s^n(\theta-\varepsilon)} \frac{\mathrm{d}b}{\Phi^n(b)}\right)\exp\left(-\int_{s^n(\theta-\varepsilon)}^{s} \frac{\mathrm{d}b}{\Phi^n(b)}\right) \\ &= 1 - [1 - P^n(\theta-\varepsilon)]\exp\left(-\int_{s^n(\theta-\varepsilon)}^{s} \frac{\mathrm{d}b}{\Phi^n(b)}\right) \end{aligned}$$

(10.82)

因为当 n 充分大时，$s^n(\theta-\varepsilon)$ 与 $P^n(\theta-\varepsilon)$ 均收敛于 0，对所有的 $b\in[s^n(\theta-\varepsilon),s]$，$\Phi^n(b)\in[\theta-\varepsilon,\theta+\varepsilon]$，因此 $P^n(\Phi^n(s))$ 的任何聚点将位于 $\{1-\exp[-s/(\theta+\varepsilon)]\}$ 与 $\{1-\exp[-s/(\theta-\varepsilon)]\}$ 之间，由于 ε 的任意性，事实上

$$P^n(\Phi^n(s))\to P(\Phi(s))=1-\exp(-s/\theta) \qquad (10.83)$$

于是，一系列不完全信息博弈的纯策略均衡“收敛于”博弈的完全信息形式的混合策略均衡。

上面我们用若干例子对混合策略均衡的再解释作详细地阐述。Harsanyi 于 1973 年证明了任何混合策略均衡可以“几乎总是”得到为给定一系列稍微扰动博弈的纯策略均衡的极限。基本思路如下：

考虑一个策略型博弈，具有有限策略集 S_i 和盈利函数 u_i。Harsanyi 以下述方式扰动盈利：令 $\theta_i(s)$ 表示范围为一个闭区间

(假设为 $[-1,1]$)的随机变量,并且令 $\epsilon>0$ 表示一个正常数(稍后它将收敛于 0)。局中人 i 的扰动盈利函数 u_i^* 依赖于他的“类型” $\theta_i=\theta_i(s)$,$(s\in S)$,且依赖于“扰动幅度” ϵ:

$$u_i^*(s,\theta_i)=u_i(s)+\epsilon\theta_i(s) \tag{10.84}$$

Harsanyi 假定局中人的类型是统计独立的。θ_i 的概率分布记作 $P_i(\cdot)$,且假定 P_i 具有关于所有 θ_i 为连续可微的密度函数 $p(\cdot)$。Harsanyi 首先证明任何局中人 i 的最佳反应基本上是唯一的纯策略。这一点在直观上相当清楚,如果 θ_i 是连续地分布的话,对于给定局中人 i 的对手的策略时,局中人 i 取两个纯策略时的盈利的一致必定是罕见事件。作为推论,在受扰动博弈中的任何均衡,$\sigma_i(\theta_i)$ 对所有 i 以及对几乎所有 $\theta=(\theta_1,\cdots,\theta_n)$ 是纯策略。Harsanyi 证明了均衡的存在,然后论证了下述重要结果:

定理 10.1 纯化定理(Purification Theorem, Harsanyi 1973):给定一组 n 个局中人和每个局中人相应的策略空间 S_i,对于 Lebesgue 测度为 1 的盈利 $\{u_i(s)\}$ 的集合,和对于$[-1,1]$上两次可微的概率密度 p_i(这些 p_i 互为独立,$i=1,2,\cdots,n$),具盈利函数 u_i 的博弈中的任何均衡是扰动盈利 u_i^* 的纯策略均衡序列当 $\epsilon\to 0$ 时的极限。更确切地说,由扰动博弈的纯策略均衡所导致的策略上概率分布收敛于未扰动博弈的均衡分布。

注意 Harsanyi 定理的陈述,有趣的是,极限博弈中的所有混合策略均衡都可以用一系列扰动博弈来“纯化”。

读者也许注意到纯化定理中关于全测度盈利集合的限制,这个似乎“过分”的条件实际上是避开某些“病态”盈利以保证定理结论的成立。存在着两个可能的问题,它们将发生“病态”盈利。第一,可能是给定的均衡只能通过全体扰动博弈的小子集中的纯策略均衡来逼近,不同的扰动博弈区分出不同的均衡。第二,弱劣策略均衡不是任何扰动博弈均衡的极限。

其实，完全信息博弈是一种理想化的博弈，因为局中人关于其对手的目标至少有一点儿不完全信息，换句话说，每个局中人不可能完完全全地知道其他局中人的目标函数。出于这种观点，我们也许可以得出一个结论——这个结论也正是 Harsanyi 所证明了的——在纯策略与混合策略之间的区别也许是人为的。

第十一章　机制设计

迄今为止，我们所做的事情是，对于给定的博弈问题，设法寻找它的均衡解。实际生活中，存在着有意义的逆向问题：给定一组 n 个局中人，在一系列可能的结局中给定他们的盈利和效用，以及他们有关这些盈利所拥有的私人信息，是否能构造一个静态 Bayes 博弈，使得该博弈的 Bayes Nash 均衡满足一定特殊的性质？例如，商业拍卖活动中拍卖形式的确定就存在这类问题。设计什么样的拍卖形式以使卖者的期望收益达到极大化正是本章将要介绍的机制设计所需要解决的问题。拍卖的形式可以有很多种，口头公开的竞拍或者用密封形式递交标价等等，因此，实际上我们要做的事是制定博弈或游戏的某些规则。先用若干例子详细阐述机制设计基本原理。

§11.1　机制设计的若干例子

例 11.1　非线性定价(nonlinear pricing)

某公司以不变的边际成本 c 生产某种产品，并卖给某消费者。消费者购买数量 q 的产品将付出价钱 T 给公司，然而这些产品将为消费者带来一定收益。设数量 q 的产品其毛盈余为 $\theta V(q)$，这里假设 V 对买卖双方是共同知识，但 θ 仅是消费者的私人信息。对共同知识 V 作下述假设是合理的：

$V(0)=0$——不发生交易时，毛盈余必为 0。

$V'>0$——随着交易量的增加，消费者的毛盈余必将随之增加。

$V''<0$—— 显然毛盈利不是交易量 q 的线性函数，由于二阶导数小于0，曲线 V 呈向上凸形状，表明随着购买量 q 的增加，毛盈利增加的速度变得越来越缓慢。

于是，消费者将接受的效用函数可以记作

$$u_1(q,T,\theta)=\theta V(q)-T \tag{11.1}$$

作为消费者的私人信息的 θ 不为公司知晓，但公司知道 θ 的概率分布：$P(\theta=\underline{\theta})=\underline{p}$，$p(\theta=\overline{\theta})=\overline{p}$，其中 $\overline{\theta}>\underline{\theta}>0$，$\underline{p}+\overline{p}=1$。博弈的进展过程如下：公司开出一个价 $T(q)$（它可能是 q 的非线性函数），如果消费者选择消费量 q 的话，他将付给公司 $T(q)$。对于这一提议，消费者要么接受，要么拒绝。一般地，可以合理地规定 $T(0)=0$—— 不发生交易时，当然不存在钱款的往来。并且我们假定会发生交易，即消费者总是接受公司的提议，这个假设当然要求公司的开价 $T(q)$ 满足约束 $\theta V(q)-T(q)\geqslant 0$，这个约束使得消费者通过交易可以有利可图，至少不会亏本。

倘若公司正确地知道 θ 的值，那么他会开价 $T(q)=\theta V(q)$，为使自己的效用极大化，公司将选择 q 使得 $\theta V(q)-cq$ 达到极大。于是这个 q 必须满足 $\theta V'(q)=c$。然而由于消费者有两个类型 $\underline{\theta}$ 与 $\overline{\theta}$，因此公司将有两套方案：$(\underline{q},\underline{T})$ 应付 $\underline{\theta}$ 类型顾客，$(\overline{q},\overline{T})$ 应付 $\overline{\theta}$ 类型顾客。于是公司的期望收益应当等于

$$Eu_0=\underline{p}(\underline{T}-c\,\underline{q})+\overline{p}(\overline{T}-c\overline{q}) \tag{11.2}$$

现在，公司面对两类约束，第一类约束是，要求顾客愿意购买，这类约束称为个人理性（IR，individual-rationality）或者参与（participation）约束。顾客不购买时的效用称之为保留效用（reservation utility），这里 $u_1(0,\theta,T(0))=\theta V(0)-T(0)=0$，$IR$ 约束用数学语言来表达，就是不管哪一种类型的顾客，他如果购买产品的话，其收益至少不应小于保留效用，即

$$(IR_1)\quad \underline{\theta}V(\underline{q})-\underline{T}\geqslant 0 \qquad (11.3)$$

$$(IR_2)\quad \bar{\theta}V(\bar{q})-\bar{T}\geqslant 0 \qquad (11.4)$$

第二类约束要求，对于每一种类型的顾客消费，公司将为该种类型设计合适的方案。这就是激励相容(incentive-compatibility)约束。用公式来表示为

$$(IC_1)\quad \underline{\theta}V(\underline{q})-\underline{T}\geqslant\underline{\theta}V(\bar{q})-\bar{T} \qquad (11.5)$$

$$(IC_2)\quad \bar{\theta}V(\bar{q})-\bar{T}\geqslant\bar{\theta}V(\underline{q})-\underline{T} \qquad (11.6)$$

现在公司必须在 IR 与 IC 两个约束条件下制定两套方案$\{(\underline{q},\underline{T}),(\bar{q},\bar{T})\}$，以使自己的期望收益极大化。或者说，在条件$(IR_1)$、$(IR_2)$、$(IC_1)$与$(IC_2)$约束下，试图求使 Eu_0 达到极大可能的$(\underline{q},\underline{T})$与$(\bar{q},\bar{T})$。为此，不妨假设$(\underline{q},\underline{T})$与$(\bar{q},\bar{T})$符合约束条件并使 Eu_0 极大化，我们来观察一下此时这 4 个约束条件之间存在着何种关系。倘若(IR_1)与(IC_2)满足，那么由(IR_1)得

$$\underline{T}\leqslant\underline{\theta}V(\underline{q})$$

故由(IC_2)得

$$\bar{\theta}V(\bar{q})-\bar{T}\geqslant\bar{\theta}V(\underline{q})-\underline{T}\geqslant\bar{\theta}V(\underline{q})-\underline{\theta}V(\underline{q})$$
$$=(\bar{\theta}-\underline{\theta})V(\underline{q})\geqslant 0 \qquad (11.7)$$

(11.7)式就是(IR_2)，显然，我们已经证明了，如果(IR_1)与(IC_2)成立的话，(IR_2)自动成立，因此 4 个约束条件可以缩减为 3 个约束条件(IR_1)、(IC_1)与(IC_2)即可。而且由(11.7)式还可以进一步看到，除非$\underline{q}=0$，否则 $V(\underline{q})>0$，且$(\bar{\theta}-\underline{\theta})>0$，因此必有 $\bar{\theta}V(\bar{q})-\bar{T}>0$，即除非公司不卖产品给低(需求)类型顾客，一般地，(IR_2)更严格地成立不等号。与此相比，(IR_1)必定成立等号，即$\underline{T}=\underline{\theta}V(\underline{q})$。因为$(IR_2)$不可能成立等号，如果$(IR_1)$也不成立等号的话，那么对于一个微小的正量 $\varepsilon>0$，令 $\bar{T}$ 代之以$(\bar{T}+\varepsilon)$，而以$(\underline{T}+\varepsilon)$代$\underline{T}$，显然这种做法一点儿也不会影响$(IC_1)$与$(IC_2)$的成立，因为

$$\underline{\theta}V(\underline{q})-(\underline{T}+\varepsilon)\geqslant\underline{\theta}V(\bar{q})-(\bar{T}+\varepsilon)$$

即为(11.5)式,而

$$\bar\theta V(\bar q)-(\bar T+\varepsilon)\geqslant\theta V(\underline q)-(\underline T+\varepsilon)$$

就是(11.6)式。而这样的$(\bar T+\varepsilon)$与$(\underline T+\varepsilon)$当然也不会影响到$(IR_1)(IR_2)$的成立,这蕴含着,如果$(IR_1)$也只成立不等号的话,那么公司的卖价不管对何种类型的顾客都可以再提高一些(至少提高 ε),这与我们关于$(\underline q,\underline T)(\bar q,\bar T)$的假设矛盾。至此,我们的约束条件应为$(IC_1)$、$(IC_2)$以及仅成立等号的$(IR_1)$。

接下来,我们将证明(IC_2)也必定只成立等号,因为如果(IC_2)可成立严格不等式的话,即

$$\bar\theta V(\bar q)-\bar T>\bar\theta V(\underline q)-\underline T \tag{11.8}$$

令

$$\varepsilon=\{[\bar\theta V(\bar q)-\bar\theta V(\underline q)+\underline T]-\bar T\}/2 \tag{11.9}$$

我们令 T 代以$(\bar T+\varepsilon)$

$$\begin{aligned}\bar\theta V(\bar q)-(\bar T+\varepsilon)&=\frac{1}{2}\{2\bar\theta V(\bar q)-\bar T-\bar\theta V(\bar q)+\bar\theta V(\underline q)-\underline T\}\\&=\frac{1}{2}\{\bar\theta V(\bar q)-\bar T+\bar\theta V(\underline q)-\underline T\}\\&>\frac{1}{2}\{\bar\theta V(\underline q)-\underline T+\bar\theta V(\underline q)-\underline T\}\\&=\theta V(\underline q)-\underline T\end{aligned} \tag{11.10}$$

即(IC_2)仍成立,而(IC_1)的成立是不需证明的,又对 $\bar T$ 增加少许量,一点也不影响只成立等号(IR_1),于是,公司还可以通过增加对 $\bar\theta$ 类型顾客的定价 $\bar T$ 而从中获益,这又与我们关于 $\bar T$ 原先的假设不符。

证明(IC_2)成立等号的方法还可以通过直观的形式,如图 11.1 所示。

对应类型$\underline\theta$的理想配置$(\underline q,\underline T)$在图中用点 A 表示、对应 $\bar\theta$ 的理想配置$(\bar q,\bar T)$在图中用点 B 表示。在图 11.1 中分别对类型 $\bar\theta$ 与$\underline\theta$消费者画出无差异曲线(indifference curve),所谓无差异曲线,即

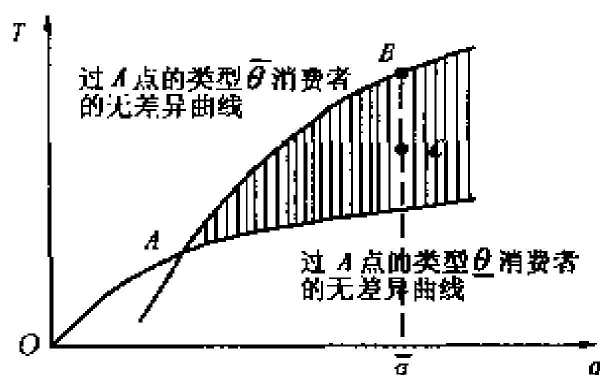


图 11.1

为 $\theta V(q)-T=c$（或 $T=\theta V(q)+c$），在经济学上它用来表示，对于某固定 θ，该类型消费者在这条曲线上的所有配置 (q,T) 的效用（或盈利）恒为常数。对类型 $\underline{\theta}$，无差异曲线方程为

$$T=\underline{\theta}V(q)+c_1=\underline{\theta}V(q) \tag{11.11}$$

$c_1=0$ 是因为 (IR_1) 成立等号，因为在上式以 $(\underline{q},\underline{T})$ 代入，并利用 (IR_1) 成立等号，立得 $c_1=0$。对于 $\bar{\theta}$ 类型的消费者的无差异曲线的方程为

$$T=\bar{\theta}V(q)+c_2 \tag{11.12}$$

其中

$$c_2=\underline{T}-\bar{\theta}V(\underline{q}) \tag{11.13}$$

此时类型 $\bar{\theta}$ 消费者的无差异曲线(11.12)式恰好通过 A 点。(11.11)式通过 A 点是自然的（因为 $c_1=0$）。

两条无差异曲线(11.11)式与(11.12)式在图 11.1 中的曲线斜率应分别为 $\underline{\theta}V'(q)$ 与 $\bar{\theta}V'(q)$，由于 $\bar{\theta}>\underline{\theta}$，因此在同一个 q 处，高类型顾客的无差异曲线总比低类型顾客的无差异曲线要陡一些。更由于这两条曲线在 A 点处相交，因而在 A 点右侧自然地高类型消费者的无差异曲线位于低类型的上方。B 点的坐标为 $(\bar{q},\bar{T})$，如果 (IC_2) 成立严格不等式，即有

$$\bar{T}<\bar{\theta}V(\bar{q})+\underline{T}-\bar{\theta}V(\underline{q})=\bar{\theta}V(\bar{q})+c_2 \tag{11.14}$$

(11.14)式表示 $B(q,T)$ 应在高类型无差异曲线相应 $\bar{q}$ 点的下方，而

$$T=\bar{\theta}V(\bar{q})\geqslant\underline{\theta}V(\bar{q})$$

蕴含着 B 点应在低类型无差异曲线的上方，由此可见，如果 (IC_2) 不成立等号，理论上 B 点应位于图 11.1 阴影部分中相应于 $\bar{q}$ 垂线上的某处（不妨记为 C），事实上，这一段证明无疑也告诉我们 $\bar{q}\geqslant\underline{q}$，即高类型消费者比低类型消费者（在均衡状态）购买更多数量产品。

我们想证明 (IC_2) 恰好只成立等号，实际上相当于证明 B 点不应当在阴影部分内，而应当位于过 C 点的垂直线与高类型消费者的无差异曲线的交叉之处。无可非议，C 点对于公司而言，绝对不是针对消费者的最佳配置，因为此时公司还可以提高 $\bar{T}$，使得可以用图中的 B 点代替 C 点，从而使公司的效用更高，既然 B 点应当位于高类型消费者的无差异曲线上，那么 (IC_2) 必定仅成立等号。

现在我们知道，求配置 $(\underline{q},\underline{T})$、$(\bar{q},\bar{T})$ 以使 EU_0 达到极大化的四个约束条件中，仅需考虑 (IR_1) 与 (IC_2) 的等号约束以及 (IC_1)，在求解过程中，不妨暂先略去 (IC_1)：如果在另外两个等号约束下求得的解也满足 (IC_1) 约束的话，我们事实上已经获得了整个问题的解。考虑约束

$$(IR_1)' \quad \underline{\theta}V(\underline{q})-\underline{T}=0 \tag{11.15}$$

$$(IC_2)' \quad \bar{\theta}V(\bar{q})-\bar{T}=\bar{\theta}V(\underline{q})-\underline{T} \tag{11.16}$$

极大化 EU_0 相当于极大化如下(11.17)式：

$$\begin{aligned}EU_0&=\underline{p}(\underline{T}-c\,\underline{q})+\bar{p}(\bar{T}-c\bar{q})\\&=\underline{p}[\underline{\theta}V(\underline{q})-c\,\underline{q}]+\bar{p}[\bar{\theta}V(\bar{q})-\bar{\theta}V(\underline{q})+\underline{\theta}V(\underline{q})-c\bar{q}]\\&=[\underline{p}\,\underline{\theta}-\bar{p}(\bar{\theta}-\underline{\theta})]V(\underline{q})-\underline{p}c\,\underline{q}+\bar{p}[\bar{\theta}V(\bar{q})-c\bar{q}]\end{aligned} \tag{11.17}$$

为求一阶条件，在(11.17)式中对 $\underline{q}$ 求偏导并使之等于 0：

$$\underline{\theta}V'(\underline{q})=c/(1-\frac{\bar{p}(\bar{\theta}-\underline{\theta})}{\underline{p}\underline{\theta}}) \tag{11.18}$$

为使上式有实际意义，我们必须假设 $\overline{p\theta}<\underline{\theta}$，它能保证分母不为 0 且为正。

再对 $\bar{q}$ 求偏导并使之等于 0，得另一个一阶条件：

$$\bar{\theta}V'(\bar{q})=c \tag{11.19}$$

高类型（从而也是高需求）消费者购买的量 $\bar{q}$ 满足，产品消耗的边际效用等于边际成本，因此它的确是令人满意且理想的，(11.19) 式满足前面所提到的方程 $\theta V'(q)=c$，即公司一旦知道类型会根据该方程定 q 与相应的 T，于是 $\bar{q}$ 也将使公司非常满意。

与此相比，低需求消费者购买的量则不那么令人满意了。若 $\underline{\theta}\neq\bar{\theta}$，从(11.18)式可知，$\underline{\theta}V'(\underline{q})>c$。即$\underline{q}$不能使 $\underline{\theta}V(q)-cq$ 达到极大值。注意到假使 q^* 是极大值点的话，应有 $V'(q^*)=c$，故有 $V'(\underline{q})>V'(q^*)$，由于 $V''<0$，即一阶导函数 $V'(\cdot)$ 单调下降，这蕴含着$\underline{q}<q^*$，就是说，公司压低了低类型消费者的购买量。这使人很容易理解，这一压低，使得高类型消费者失去购买$\underline{q}$量的兴趣（或者放弃声称购买$\underline{q}$量以进行“欺诈”的想法）。高类型消费者仍有兴趣表示自己的高需求，则使得公司可以提高 $\bar{T}$，或者，等价地，由于 $(IC_2)^*$，缩减高需求消费者的经济收益$(\bar{\theta}-\underline{\theta})V(\underline{q})$（根据 $(IC_2)^*$，$(\bar{\theta}-\underline{\theta})V(\underline{q})=\bar{\theta}V(\bar{q})-\bar{T}$ 恰为高需求顾客的效用）。于是，公司为了提高收益这一目的，最好的办法是牺牲某些效益。进一步，如果$(IR_1)^*$与$(IC_2)^*$成立，那么，费用 $\bar{T}$ 与$\underline{T}$可以用 $\bar{q}$ 与$\underline{q}$来确定。

最后，我们要检验一下，上述得到的解是否也使 IC_1 满足，这是件很要紧的事情。也就是说，我们要检验

$$\underline{\theta}V(\underline{q})-\underline{T}=0\geqslant\underline{\theta}V(\bar{q})-\bar{T}$$

由于$(IC_2)^*$的成立，我们有

$$\underline{\theta}V(\bar{q})-\bar{T}=\underline{\theta}V(\bar{q})-\bar{\theta}V(\bar{q})-(\bar{\theta}-\underline{\theta})V(\underline{q})$$

$$= -(\bar{\theta}-\underline{\theta})(V(\bar{q})-V(\underline{q}))\leqslant 0 \qquad (11.20)$$

而 IR_1 的成立必使 IR_2 也成立这一点已在前面论证过。现在我们可以说上述办法得到的解 $(\underline{q},\underline{T})(\bar{q},\bar{T})$ 应为全部问题之解。

例 11.2　拍卖

假设卖者有一单元货物待售，有两个可能买者 $(i=1,2)$，他们可视作事前恒同。各人对于该货物的估价 θ_1 与 θ_2，均以概率 $\underline{p}$ 取值 $\underline{\theta}$ 而以概率 $\bar{p}$ 取值 $\bar{\theta}$，其中 $\underline{p}+\bar{p}=1$，且二人的估价看作是相互独立的变量。每一个买者当然知道自己的估价，但卖者与另一个买者却不知道。

卖者可以推出各种不同的拍卖形式，哪一种拍卖形式使卖者的盈利达到极大化呢？为了解答这个问题，我们试图求解卖者的最优机制。

假设卖者在两个买者之间建立某些"信号博弈"，即建立一些发送与接受信号的规则。并且说明货物的配给与费用依赖于所选择的信号。以 s_1 与 s_2 分别表示在博弈中两买者策略 σ_1 与 σ_2 的具体实施。所谓机制，必须确定货物转让给买者 i 的概率 $x_i(s_1,s_2)$ 以及为此买者 i 向卖者支付的费用 $T_i(s_1,s_2)$。

令 $\{\sigma_1^*(\cdot),\sigma_2^*(\cdot)\}$ 表示博弈或所设计的机制中的 Bayes 均衡策略。因为买者可以自由地不去参与拍卖，因此，买者 1 的个人理性约束（即 IR）应为，对于每一个 θ_1 和构成 $\sigma_1^*(\theta_1)$ 的支撑集的每一个 s_1，成立

$$(IR_1)\quad E_{\theta_2}E_{\sigma_2^*(\theta_2)}[\theta_1 x_1(s_1,s_2)-T_1(s_1,s_2)]\geqslant 0 \quad (11.21)$$

同样地，买者 2 的 IR 约束为

$$(IR_2)\quad E_{\theta_1}E_{\sigma_1^*(\theta_1)}[\theta_2 x_2(s_1,s_2)-T_2(s_1,s_2)]\geqslant 0 \quad (11.22)$$

这表明不管买者得到货物与否，他的效用至少不会比他不参与拍卖时的保留效用小。至于买者 1 的 IC 约束，则是对每一个 θ_1，相应于 $\sigma_1^*(\theta_1)$ 的支撑集的每一个 s_1，以及每一个其他的 s_1'，

$$(IC_1)\quad E_{\theta_2}E_{\sigma_2^*(\theta_2)}[\theta_1 x_1(s_1,s_2)-T_1(s_1,s_2)]$$
$$\geqslant E_{\theta_2}E_{\sigma_2^*(\theta_2)}[\theta_1 x_1(s_1',s_2)-T_1(s_1',s_2)] \quad (11.23)$$

类似地得到买者 2 的 IC 约束条件

$$(IC_2)\quad E_{\theta_1}E_{\sigma_1^*(\theta_1)}[\theta_2 x_2(s_1,s_2)-T_2(s_1,s_2)]$$
$$\geqslant E_{\theta_1}E_{\sigma_1^*(\theta_1)}[\theta_2 x_2(s_1,s_2')-T_2(s_1,s_2')] \quad (11.24)$$

上面四个约束条件需要说明的是盈利或效用函数的定义，以买者 1 为例，他对货物的估计是 θ_1，若他获得此物的概率为 $x_1(s_1,s_2)$，那么他的期望收益当为 $\theta_1 x_1(s_1,s_2)$，$T_1(s_1,s_2)$表示他为此支付的费用，当然，如果他所取的 s_1 表示不买，或他的投标小于买者 2 而得不到货物，此时可定义 $T_1(s_1,s_2)=0$。因此，从拍卖中买者 1 获得的盈利当为 $\theta_1 x_1(s_1,s_2)-T_1(s_1,s_2)$。但是这种计算是基于买者 2 的策略行动 s_2 为前提，而 s_2 是买者 2 的均衡策略 $\sigma_2^*(\theta_2)$的一个支撑点，于是买者 1 的盈利应对 $\sigma_2^*(\theta_2)$求期望，而类型 θ_2 有其分布，从而必须对 θ_2 再一次地求期望。

读者不难发现，倘若不得不考虑所有可能的信号空间的话，卖者很难确定最佳拍卖机制。幸好人们可以把注意力限于"直接显示博弈"(direct-revelation games)，在直接显示博弈中，两个买者同时报告(可能不太诚实)他们的类型$(\hat{\theta}_1,\hat{\theta}_2)$。现定义两个买者的消费函数及费用函数：

$$\tilde{x}_i(\hat{\theta}_1,\hat{\theta}_2)=E_{(\sigma_1^*(\hat{\theta}_1),\sigma_2^*(\hat{\theta}_2))}[x_i(s_1,s_2)] \quad (i=1,2) \quad (11.25)$$

$$\tilde{T}_i(\hat{\theta}_1,\hat{\theta}_2)=E_{(\sigma_1^*(\hat{\theta}_1),\sigma_2^*(\hat{\theta}_2))}[T_i(s_1,s_2)] \quad (i=1,2) \quad (11.26)$$

注意(11.25)式和(11.26)式是基于买者所报告的类型所相应的均衡策略 σ_1^* 与 σ_2^* 求期望的。也就是说，$\tilde{x}_i$、$\tilde{T}_i$ 是直接显示博弈相应的 x_i、T_i。我们的约束条件 IR 保证了买者愿意参与直接显示博弈，参与拍卖既然不至于吃亏，那么希望得到货物的买者自然是愿意参与的。事实上，最优设计的机制或博弈都必须保证博弈局中人的自愿参与，这是博弈的先决条件，因此机制设计中 IR 约束条件几乎是第一位的，就像上市公司推出自己的股票时，每股溢价不

可能漫天要价，一旦股民发现购买这类股票几乎肯定被套，那么没有人愿意“上当”，从而博弈无法进行下去，发行股票也就相当于失败。当然溢价也不可能相当地低，否则卖者的盈利谈不上极大化，这是机制设计紧接着要解决的问题。至于约束条件 IC，则保证了直接显示博弈中 Bayes 均衡将要求两个买者都讲老实话，也就是 $\hat{\theta}_1=\theta_1,\hat{\theta}_2=\theta_2$。我们将在下面详细论述。

现在我们来求解最优对称拍卖，稍后我们也将会看到，最佳拍卖的确是对称的。注意到四个约束条件 (IR_1)、(IR_2)、(IC_1) 与 (IC_2)，仅仅包含了每一位买者得到货物的期望概率以及支付给卖者的期望费用，而且期望是关于另一个买者的类型而求得的，因此，不妨令 $\bar{x}$、$\underline{x}$、$\bar{T}$、$\underline{T}$ 分别表示具有类型 $\bar{\theta}$ 与 $\underline{\theta}$ 的买者获得货物的期望概率与期望费用。由于考虑的是对称均衡，因此这些记号中无需下标 i。以这些记号重新罗列约束条件 IR 与 IC：

$$(IR_1)\quad \underline{\theta}\,\underline{x}-\underline{T}\geqslant 0 \tag{11.27}$$

$$(IR_2)\quad \bar{\theta}\,\bar{x}-\bar{T}\geqslant 0 \tag{11.28}$$

$$(IC_1)\quad \underline{\theta}\,\underline{x}-\underline{T}\geqslant \underline{\theta}\bar{x}-\bar{T} \tag{11.29}$$

$$(IC_2)\quad \bar{\theta}\,\bar{x}-\bar{T}\geqslant \bar{\theta}\,\underline{x}-\underline{T} \tag{11.30}$$

读者不难发现，一旦使用了期望概率与期望费用的记号，我们将本例中约束条件的形式简化得与例 11.1 非线性定价时一样。假如卖者卖出货物的机会成本由于与货物之值相比微不足道而被忽略为 0，不管货物转给哪一个买者，卖者的期望收益应等于

$$Eu_0=\underline{p}\,\underline{T}+\bar{p}\bar{T} \tag{11.31}$$

完全与例 11.1 讨论约束条件的过程一样，只有 (IR_1) 与 (IC_2) 恰好成立等号，而其他两种约束条件 (IR_2)、(IC_1) 不成立等号。于是，(IR_1) 与 (IC_2) 一起确定了期望费用：$\underline{T}=\underline{\theta}\,\underline{x}$ 与 $\bar{T}=\bar{\theta}(\bar{x}-\underline{x})+\underline{\theta}\,\underline{x}$。将它们代入 Eu_0 得

$$Eu_0=(\underline{\theta}-\bar{p}\bar{\theta})\underline{x}+\bar{p}\bar{\theta}\bar{x} \tag{11.32}$$

迄今为止，我们尚未对期望概率 $\underline{x}$ 与 $\bar{x}$ 强加过任何约束。倘若买者

只有一个，毫无疑问，$0\leqslant\underline{x},\overline{x}\leqslant 1$。但如果存在两个买者，不能不面对这样的现实可能：一个买者获得货物，与此同时，另一个买者必定得不到货物。起码，在知道买者的类型之前，买者得到货物的概率不会超过$\frac{1}{2}$：

$$\underline{p}\ \underline{x}+\overline{p}\ \overline{x}\leqslant\frac{1}{2} \tag{11.33}$$

(11.33)式由对称性而得，按理，在不知道任何一个买者的类型之前，每个买者得到货物的概率各占一半，为$\frac{1}{2}$，但由于存在着货物没有拍卖成功的可能，因而二买者得到货物的事前概率应为小于等于1，于是自然成立(11.33)式。(11.33)式实际上构成了关于$\underline{x}$与$\overline{x}$的约束条件。

首先，假设$\underline{\theta}\leqslant\overline{p}\theta$。观察(11.32)式，$\underline{x}$前的系数为非正，于是$Eu_0$随$\underline{x}$的增加而非增，而由于$\overline{p}\theta>0$，$Eu_0$随$\overline{x}$的增加而增加。为使$Eu_0$极大化，卖者自然希望$\underline{x}=0$且$\overline{x}$"尽可能地大"。但是，$\underline{x}$可以低到取0，$\overline{x}$却不能在合理范围[0,1]内"肆无忌惮"地大，因为$\overline{x}$表示的是高估价买者获得货物的期望概率。如前所述，该期望是关于另一个买者的类型求出的，那位买者的类型可能为低(概率为$\underline{p}$)，也可能为高(概率为$\overline{p}$)，然而，当两位买者都具高估价时，他们之间只有一位能获得货物，此时获得货物的条件概率应为$\frac{1}{2}$。据上述分析，$\overline{x}$不可能超过$\underline{p}+\frac{\overline{p}}{2}$。既然卖者希望$\overline{x}$尽可能地大，故至多取$\overline{x}=\underline{p}+\frac{\overline{p}}{2}$。综合这些知识，卖者为极大化自己的收益，在$\underline{\theta}\leqslant\overline{p}\theta$的时候，他设计的最优拍卖机制是：$\underline{x}=0$，这蕴含着如果两买者均宣布(标出)低估价$\underline{\theta}$，卖者拒绝卖货给他们；如果一个买者宣布$\underline{\theta}$，另一个宣布$\overline{\theta}$，那么卖者出售货物给高估价买者；又假如二买者都宣布θ，那么以$\frac{1}{2}$概率随便卖给其中的一个人。

其次，假设$\underline{\theta}>\overline{p}\overline{\theta}$，显然，$Eu_0$ 为$\underline{x}$与 $\overline{x}$ 两变量的递增函数，卖者为使收益极大化，不会取$\underline{x}=0$，换句话说，卖者不会将货物留下来不卖，于是(11.33)式必定只成立等号。$\underline{p}\,\underline{x}+\overline{p}\,\overline{x}=\frac{1}{2}$，或$\underline{x}=(\frac{1}{2}-\overline{p}\,\overline{x})/\underline{p}$，代入(11.32)式得到(利用$\underline{p}+\overline{p}=1$)

$$Eu_0=\frac{1}{2\underline{p}}(\underline{\theta}-\overline{p}\overline{\theta})+\frac{\overline{p}}{\underline{p}}(\overline{\theta}-\underline{\theta})\overline{x} \qquad (11.34)$$

(11.34)式指出 Eu_0 是 $\overline{x}$ 的递增函数，因此 $\overline{x}$ 应越大越好，利用$\underline{\theta}\leqslant\overline{p}\overline{\theta}$ 情况时的分析，取 $\overline{x}=\underline{p}+\overline{p}/2$。再以此 $\overline{x}$ 代回(11.33)式得$\underline{x}=\underline{p}/2$。解得的 $\overline{x}$ 与$\underline{x}$将产生卖者的最优拍卖机制：$\overline{x}$ 仍蕴含了只有一人宣布高估价时则由他获得货物，若二人宣布高估价时，各人以$\frac{1}{2}$概率获货物，而$\underline{x}=\underline{p}/2$ 则蕴含了两买者都宣布低估价，那么卖者以$\frac{1}{2}$概率随便卖给其中一人。

§11.2 显示准则与机制设计

上一节的两个例子初步刻画了机制设计的基本内容，以例 11.2 为例，它实际上是不完全信息博弈，在此类博弈中，总存在着一个“委托人”(principal)，他乐意把自己的行动限制在某些为其他局中人私人所知的信息的条件下，那些其他局中人称作“代理人”(agents)，委托人可以要求代理人提供个人信息，然而这些代理人可能不说真话，除非委托人给予激励以促使他们讲实话。在例 11.2 中，卖者是委托人，两个买者的身份是其他局中人或代理人。

机制设计方法的显著特点是，假定委托人为了使得自己的期望效用达到极大化，选择了与因历史或惯例原因所使用的特定机制不相同的一个机制。

机制设计的许多应用考虑拥有单个代理人的博弈，此类模型

也可应用于代理人具有连续统的情况，每一个代理人与委托人互相影响，但不与其他代理人互相影响。例 11.1 就是单个代理人情况的机制设计，公司是委托人，消费者为单个代理人，委托人设计一个机制（最佳非线性定价方案）使自己期望效用极大化。以现实经济生活为例，最佳税制研究也适合此类博弈。委托人——政府，乐于从代理人（消费者）那儿提高税收以充实公共财产。征税的最佳水平依赖于消费者的挣钱能力。如果政府知道该消费者的挣钱能力，就可以基于该能力向他征收既令政府满意又不至于与消费者的劳动量相悖的税额。在有关消费者的挣钱能力方面出现不完全信息的情况下，政府只能在他的现实收入基础上征收收入税。收入税一览表可以看作引出有关消费者能力的信息的激励方案。

机制设计也可应用于若干代理人的博弈。例 11.2 中，委托人（卖者）就是面对两个买者（代理人）设计拍卖机制，他不知道买者为买此货愿意支付多少，因此设立一个机制来确定由哪一个买者购买并确定出售价格，这个机制的设计是委托人将自己的行动限制于一定的约束且使买者知道，但最终目的是使委托人自己达到极大期望收益。

机制设计通常以三步不完全信息博弈进行研究，在博弈中，代理人的类型——例如在拍卖机制设计中买者愿支付多少——属于私人信息。

第一步：委托人设计一个“机制”（或契约，或激励方案），实际上是一个博弈规则，其中代理人发出无需成本的信号，机制也包含了依赖于已发出信号的“配给”。在例 11.2 中，卖者设计的机制，要求代理人报告自己的类型（估价多少），基于这些信号，机制设计了配给方案：货物以什么样的价格出售给高类型代理人，信号博弈可以是同时发布信号，也可以是一个更为复杂的通讯过程。

第二步：代理人同时接受或拒绝委托人所设计的机制。拒绝就是等于不参加机制或博弈，拒绝者得到某种额定的“保留效用”

(reservation utility)(通常，但未必一定，它是一个与类型无关的数)。在我们研究的机制设计中，常常假定代理人总是接受机制，因为我们"规定"委托人具有约束条件 IR，代理人参与博弈至少可得到保留效用。

第三步：接受机制的代理人进行由机制所确定的博弈。

根据机制设计的例解以及上面一段描述，现在我们正式叙述它的一般形式。

设有$(n+1)$个局中人，其中一个为委托人(不妨记作局中人0)，他没有私人信息，另外n个代理人(局中人$i=1,2,\cdots,n$)，具有某集合Θ上的类型$\theta=(\theta_1,\cdots,\theta_n)$。$\Theta$上的概率分布使得下面所要提到的效用函数存在期望与条件期望。

已知机制设计的目的在于委托人要想确定一个于己有利的合理的分配(allocation)，记作向量$y=\{x,t\}$，其中x向量称为决策(decision)，x属于一个紧且凸的非空集合$X\subset R^n$(集合X为非空这一点最易为大家所接受，其实另外两个要求也是合理的，假如x与x^*均为决策，我们自然要求他们的凸组合也是决策，这就要求集合为凸；而要求一连串决策的聚点仍为决策是相当合理的，从而要求集合具有紧致性)。分配的另一部分t则是从委托人到代理人的货币转让向量$t=(t_1,\cdots,t_n)$(由于t的定向规定是从委托人到代理人，因此它可以为正也可以为负)。一般情况下，即大多数应用中，总是取决策空间X足够地大，以使得确保有内点解，可惜，我们上面的拍卖例子恰好是个例外。

局中人$i(i=0,1,\cdots,n,)$具有效用函数$U_i(y,\theta)$。由于t的方向性，我们假定对于$i=1,2,\cdots,n$，U_i关于t_i严格地增加，即从委托人那儿转向局中人i的货币量越多，该局中人的效用将越大。与此相反，U_0当然关于每一个t_i均为递减函数。为讨论的方便起见，$U_i(i=0,1,\cdots,n)$均假设为两次可微函数。

给定了一个(依赖于类型的)分配$\{y(\theta)\}_{\theta\in\Theta}$,局中人$i(i=1,\cdots,n)$,或代理人$i$若具有类型$\theta_i$,那么它具有期望效用:

$$U_i(\theta_i)=E_{\theta_{-i}}[u_i(y(\theta_i,\theta_{-i}),\theta_i,\theta_{-i})|\theta_i] \qquad (11.35)$$

并且委托人具有期望效用

$$E_\theta u_0(y(\theta),\theta) \qquad (11.36)$$

在我们所讨论的机制设计应用中,若不作另外的声明,那么总是设代理人i的效用依赖于他自己的费用t_i和类型θ,不依赖于t_{-i}与θ_{-i}。

读者在具体应用中,应知道上述一般模型中的记号与自己所处理模型中的变量的相对应性。例如,在前面提到的最佳税收机制设计中,代理人(即消费者)的收人是x,根据x来确定y——税收标准。代理人应负的税则为t,至于θ则是代理人的类型——他的挣钱能力,这是私人信息,不为委托人(政府)所知道。

一个设计好的机制为每一个代理人i确定了他的信号空间M_i,并确定了该局中人宣布信号的博弈形式。由于类型是私人信息,y可以仅通过代理人的信号而依赖于θ,我们用y_m表示这个函数,$y_m:\mu=(\mu_1,\cdots,\mu_n)\to Y=XXR^n$,即$y_m(\mu(\theta))$。

至此,我们全部完成了机制设计模型的建立。由于机制设计的博弈存在若干阶段,我们知道在完全信息的多阶段博弈中,追求的是子博弈完美均衡,因为它与Nash均衡之间存在的差异是众所周知的。于是,这个事实提醒了我们,在机制设计博弈中寻求Bayes均衡也有存在缺陷的可能。幸好,我们已经指出我们所研究的机制设计限于在第二步中所有的代理人都接受委托人在第一步所设计的机制,这相当于不必再考虑第二步的各种结果。此外,我们像例11.2那样采用“直接显示博弈”,即让代理人同时报告自己的类型,也就是将各代理人的信号空间缩小为他们的类型空间。在下面将介绍的“显示准则”(revelation principle)作为一个简单且基本、同时又非常重要的结果证明了:为了使自己的期望收益达到

最高，委托人非但可以限于"省略"第二步，而且还可以限于在第三步中所有代理人同时讲实话的情况，就是说，所有代理人都真实地展示他们的类型。不难看出，这两个"限制"使得讨论的范围大大地缩小，这正是显示准则最奥妙也最有用的地方：委托人仅需通过代理人之间的静态 Bayes 博弈就可以得到自己最大的期望收益。

实际上，在机制设计的第三步中与第一步委托人所设计的机制相关联的博弈形式，加上第二步中代理人的接受决策一起，由于信号空间 M_i 包含类型空间，因此构成了比直接显示博弈更大一些的博弈，我们不妨称它为原博弈。不失一般性，我们可以在原博弈将各代理人的"接受机制"决策也包含进信号空间 $Mi(\cdot)$中去。现在考虑原博弈的 Bayes 均衡，为记号的方便起见，假定 Bayes 均衡是纯策略均衡，我们可以将它记作$\{\mu_i^*(\theta_i)\}(i=1,2,\cdots,n)$。

现在构造直接显示博弈，即对每一个代理人 i 有一个新的信号空间 Θ_i，它即为类型空间，因此 $\Theta_i\subseteq M_i$。每一个代理人 i 宣布一个类型 $\hat{\theta}_i\in\Theta_i$，$\hat{\theta}_i$ 未必一定等于真值 θ_i。令 $\hat{\theta}=(\hat{\theta}_1,\cdots,\hat{\theta}_n)$，定义直接显示博弈中新的分配规则 $\bar{y}$，显然 $\bar{y}$ 应是从 $\Theta=\underset{i}{X}\Theta_i$ 到 Y 上的映照，形式为

$$\bar{y}(\hat{\theta})=y_m(\mu^*(\hat{\theta}))\tag{11.37}$$

定义(11.37)式指出，在直接显示博弈中，依据各代理人宣布的类型 $\hat{\theta}$ 的分配相当于在原博弈中类型取 $\hat{\theta}$ 时 Bayes 均衡分配。这里，$\mu^*(\hat{\theta})=(\mu_1^*(\hat{\theta}_1),\cdots,\mu_n^*(\hat{\theta}_n))$。

我们将指出，在给定$\{\mu_i^*\}$是原博弈的 Bayes 均衡条件下，各个代理人讲实话：$\{\hat{\theta}_i=\theta_i\}$是直接显示博弈的 Bayes 均衡。也就是说，在其他代理人说实话时，代理人 i 只有说实话才能使自己的期望效用达到极大。

对所有 i 与 θ_i，由(11.37)式可得

$$E_{\theta_{-i}}[u_i(\bar{y}(\theta),\theta_i,\theta_{-i})|\theta_i]$$

$$=E_{\theta_{-i}}\{u_i(y_m(\mu^*(\theta)),\theta_i,\theta_{-i})|\theta_i] \qquad (11.38)$$

由于 $\mu_i^*(\theta_i)$是原博弈的 Bayes 均衡策略，因此在其他代理人分别取 $\mu_1^*(\theta_1),\cdots,\mu_{i-1}^*(\theta_{i-1}),\mu_{i+1}^*(\theta_{i+1}),\cdots,\mu_n^*(\theta_n)$时，对于 M_i 中所有的 μ_i 应当有

$$E_{\theta_{-i}}[u_i(y_m(\mu^*(\theta)),\theta_i,\theta_{-i})|\theta_i]$$
$$=\sup_{\mu_i\in M_i}E_{\theta_{-i}}[\mu(y_m(\mu_1^*(\theta_1),\cdots,\mu_i,\cdots,\mu_n^*(\theta_n)),\theta_i,\theta_{-i})|\theta_i] \qquad (11.39)$$

(11.39)式右端取上确界时 μ_i 的范围是代理人 i 的信号空间 M_i，如果 μ_i 的所属范围缩小，例如为，$\mu_i\in\{\mu_i^*(\hat{\theta}_i)\}_{\hat{\theta}_i\in\Theta_i}\subset M_i$，那么在取上确界的话，其上确界可能不会大于 $\mu_i\in M_i$ 时所取的上确界，用数学公式表示，则为

$$\sup_{\mu_i\in M_i}E_{\theta_{-i}}[u_i(y_m(\mu_1^*(\theta_1),\cdots,\mu_i,\cdots,\mu_n^*(\theta_n)),\theta_i,\theta_{-i})|\theta_i]$$
$$\geqslant\sup_{\mu_i\in\{\mu_i^*(\hat{\theta}_i)\}_{\hat{\theta}_i\in\Theta_i}}E_{\theta_{-i}}[u_i(y_m(\mu_1^*(\theta_1),\cdots,\mu_i,\cdots,\mu_n^*(\theta_n)),\theta_i,\theta_{-i}|\theta_i]$$
$$=\sup_{\hat{\theta}\in\Theta_i}E_{\theta_{-i}}[u_i(y_m(\mu_1^*(\theta_1),\cdots,\mu_i^*(\hat{\theta}_i),\cdots,\mu_n^*(\theta_n)),\theta_i,\theta_{-i}|\theta_i]$$
$$=\sup_{\hat{\theta}_i\in\Theta_i}E_{\theta_{-i}}[u_i(\overline{y}(\theta_1,\cdots,\hat{\theta}_i,\cdots,\theta_n),\theta_i,\theta_{-i})|\theta_i] \qquad (11.40)$$

(11.40)式中最后一个等号又是由于 $\overline{y}$ 的定义所致。综合(11.38)式、(11.39)式与(11.40)式。我们得到

$$E_{\theta_{-i}}[u_i(\overline{y}(\theta),\theta_i,\theta_{-i})|\theta_i]$$
$$\geqslant\sup_{\hat{\theta}_i\in\Theta_i}E_{\theta_{-i}}[u_i(\overline{y}(\theta_1,\cdots,\hat{\theta}_i,\cdots,\theta_n),\theta_i,\theta_{-i})|\theta_i] \qquad (11.41)$$

(11.41)式最明白无误地说明了对于代理人 i 来说，如果其他代理人讲真话，那么他最好也讲真话。

如果考虑的不是原博弈的纯策略均衡，而考虑均衡 σ^* 是混合策略的，那么定义 $\overline{y}$ 为 $\hat{\theta}$ 的适当的随机函数，上述推理照样成立，此处不再赘述。

我们将上面的有关论证总结为定理形式：

定理 11.1 显示准则(revelation principle):假如一个具有信号空间 $M_i(i=1,2,\cdots,n)$和分配函数 $y_m(\cdot)$的机制有 Bayes 均衡

$$\mu^*(\cdot)=\{\mu_i^*(\theta_i)\}_{i=1,2,\cdots,n;\theta_i\in\Theta_i}$$

其中 Θ_i 是代理人 i 的类型空间。那么,存在一个直接显示机制(即 $\bar{y}=y_m(\mu^*(\cdot))$),它的信号空间恰为类型空间,并且该直接显示博弈存在一个 Bayes 均衡,其中所有代理人都在机制设计的第二步接受机制,且在第三步博弈中真实地报告他们各自的类型。

显示准则揭示了这样一个事实:任何一个机制所能达到的 Bayes 均衡分配结果都可以通过一个讲真话的直接机制来实现。因此,委托人可以只考虑直接机制设计。

机制设计问题是博弈理论中的一个重要课题,它包含了十分丰富的内容,例如代理人为单个或多个的情况。由于数学内容的复杂性,本书不准备详细介绍。有兴趣的读者可以阅读 Fudenberg & Tirole 的《Game Theory》中的 Ch. 7。

第五部分　不完全信息动态博弈

第十二章　完美 Bayes 均衡

众所周知，在完全信息博弈中，子博弈完美均衡实质上是对 Nash 均衡的精炼。因为在展开型博弈中，某些 Nash 均衡由于空头或不可信的威胁与承诺而不宜作为博弈的合理预测，子博弈完美正是通过在每一个子博弈中要求均衡策略剖面的实现都是 Nash 均衡，从而排除和摈弃了这类不合理预测。如果博弈属于不完全信息，子博弈完美是否也可以对 Bayes Nash 均衡同样地进行精炼呢？答案是令人失望的。在不完全信息博弈中，即使每个局中人观察到其他局中人在每个周期结束时的行动，由于他不知道对方所属类型，在即将开始的周期内无法确定将要进行的是哪一个子博弈，于是我们不能够检验后续策略是否为 Nash 均衡。我们知道，局中人属于何种类型被假设为由“自然”赋予的概率而确定，所谓信念实际上是先验信念。要像子博弈完美那样地精炼 Nash 均衡，在不完全信息动态博弈中需要局中人在观察到上一周期的行动之后“适当地修正”关于对手类型的信念——在统计学中称之为后验信念，显然这种后验信念可以通过概率统计中的 Bayes 法则得到解决。本章介绍的完美 Bayes 均衡正是利用后验信念的获得

而使具行动的局中人为自己确定最优策略，从而达到精炼 Bayes Nash 均衡的目的。因此，完美 Bayes 均衡是对 Bayes 均衡的一种精炼，又是子博弈完美思想在不完全信息博弈中的伸展，它本身当然是 Nash 均衡。

§12.1 完美 Bayes 均衡定义

为引进完美 Bayes 均衡这个重要概念，首先考虑一个完全但不完美信息的动态博弈，见图 12.1。

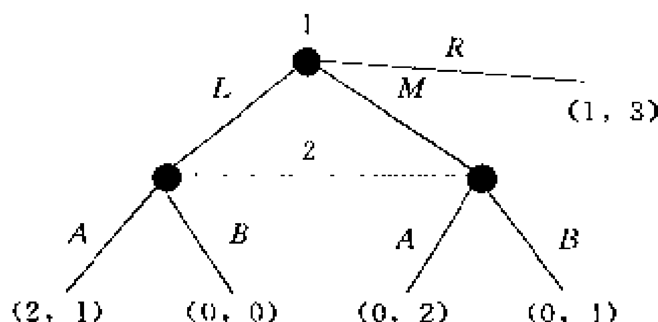


图12.1

首先，由局中人 1 在 L、M、R 三个行动中进行选择，如果他选择了 R，则博弈结束。如果局中人 2 知道局中人 1 没有选择 R，他必然知道局中人 1 要么选择 L 要么选择 M。但图 12.1 告诉我们，局中人 2 不知道局中人 1 究竟在 L 或 M 中选择了哪一个行动——因此我们称局中人信息是完全但并不完美，在不完美信息下他开始选择自己的行动，要么选 A，要么选 B，然后结束博弈。各个可能结局的相应盈利向量列在图 12.1 的终点结处。为分析该博弈，根据图 12.1 不难给出相应的正则型表示，见图 12.2。易见该博弈有两个纯策略 Nash 均衡：(L,A)与(R,B)。由于这是一个动态博弈，我们关心这些策略剖面是否子博弈完美均衡。从图 12.1 可知，除了原来的博弈之外，该博弈不存在任何其他真子博弈，子博弈完美的要求将自然且平凡地满足。于是，在此类展开型博弈

中，子博弈完美 Nash 均衡定义实际上就是 Nash 均衡。结论不言而喻，(L,A)与(R,B)都是子博弈完美均衡。

		局中人 2	
		A	B
	L	2,1	0,0
局中人 1	M	0,2	0,1
	R	1,3	1,3

图 12.2

然而，事情并未圆满结束。直觉上，(R,B)显然依赖于一个不可信威胁：对于局中人 2 来说，如果他开始行动，无疑 B 是个劣策略，B 至少弱劣于 A。正因为如此，局中人 1 决不会(只要他是理性的)因为局中人 2 威胁“取 B”而被迫采取策略 R，为极大化自己的盈利，局中人 1 会取 L，迫使局中人 2 只得采取行动 A。概括一句话，局中人 1 认为“局中人 2 将取 B”只不过是个空头威胁。

上述例子反映出一个事实，在信息完全但不完美的博弈中，尽管(R,B)是子博弈完美的，可是它仍依赖于一个不可信的空头威胁。因此它理应从我们的合理预测中排除出去。倘若博弈模型更趋复杂一些，对均衡的进一步精炼看来完全有必要，目的在于剔除“不可置信”即“不合理”均衡，因此我们必须对均衡“强加”一些条件。

上例出现的主要问题之一是，局中人 2 在他将要采取行动的信息集上不知道哪一个结将会达到，也就是他完全不知道，局中人 1 若不取 R，那么他究竟取 L 还是取 M。我们强加的条件中，将要求局中人 2 对于这个问题要有一定的信念，并在这个信念下采取最佳的策略行动，这就是 R_1 与 R_2 条件：

R_1：在每一个信息集，在该集具有行动的局中人关于博弈到达信息集中的哪一个结必须有一个信念。对于非单结信息集，信念是

信息集中各个结上的概率分布。至于单结信息集，信念则置概率 1 于单决策结上。

R_2：在给定的信念下，局中人的策略必须是序贯理性的(sequentially rational)。就是说，在每一个信息集，具行动的局中人所采取的行动(以及局中人往后的行动)在给定该局中人在该信息集上的信念与其他局中人以后的策略下必须是最优的。

为使我们的例子满足 R_1 与 R_2，现在在局中人 2"不完美"的信息集上赋予一个概率分布(p,1－p)作为信念，重新绘制博弈树如图 12.3。

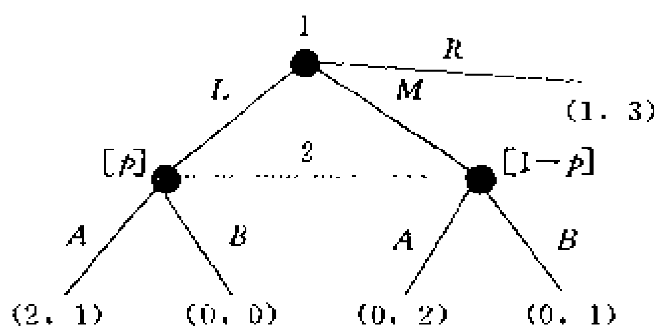


图12.3

一旦给定局中人 2 在"不完美"信息集上的信念，可以计算局中人 2 的期望盈利。若局中人 2 取 A，期望盈利为

$$p\cdot 1+2(1-p)=2-p \tag{12.1}$$

若局中人 2 取行动 B，相应的期望盈利为

$$p\cdot 0+(1-p)\cdot 1=1-p \tag{12.2}$$

无论 p 取何值，只要 $0\leqslant p\leqslant 1$，总有 $2-p>1-p$，根据 R_2 要求，局中人 2 不应取 B。我们发现，对图 12.1 展开型博弈赋予要求 R_1 与 R_2 之后，事实上已经排除了前面提及的那个依赖于空头威胁从而不怎么合理的子博弈完美均衡(R,B)。

R_1 与 R_2 的作用已经显示出来，作为对均衡精炼的要求，R_1 与 R_2 坚持局中人必须有信念并且在这些信念下采取最优行动。可是他们没有对局中人的信念提出任何要求。最起码地，我们设想

"所赋予的"信念应当是合理的，只有合理的信念才可能有合理的预测。为了对局中人的信念进一步强加要求"它是合理的"，我们必须区分均衡路径上的信息集与非均衡路径上的信息集。首先要求弄清什么是均衡路径，这个概念以前曾在介绍展开型博弈时模糊地涉及，现在给予清晰的说明。

给定展开型博弈中的一个均衡(无论是前面介绍过的 Nash 均衡、子博弈完美均衡、Bayes Nash 均衡，还是稍后就要介绍的完美 Bayes 均衡)，如果按照均衡策略进行博弈，将以正概率到达的信息集，则称这个信息集在均衡路径上，如果按照均衡策略肯定达不到的信息集，则被称作不在均衡路径上(off the equilibrium path)。

根据定义，均衡路径显然与均衡策略剖面有关，不同的均衡策略剖面有不同的均衡路径。

现在先对均衡路径上的信念作出要求：

R_3：局中人在均衡路径上信息集的信念，是通过 Bayes 法则与局中人的均衡策略来确定的。

考虑图 12.3 的子博弈完美均衡(L,A)中，给定了局中人 1 的均衡策略 L，根据前面的计算，局中人 2 知道在不完美的信息集中到达的是哪一个结，显然局中人 2 的信念必定是 $p=1$。倘若换了另一个均衡，譬如说该模型中存在另外一个混合策略(注：这里仅仅是假设，是否真的存在并不影响我们对问题的解释)；局中人 1 以概率 q_1 取 L，以概率 q_2 取 M，以概率$(1-q_1-q_2)$取 R，利用 Bayes 法则，不难得到局中人 2 的信念：$p=q_1/(q_1+q_2)$。

对于我们即将介绍的新概念——完美 Bayes 均衡来说，R_1－R_3 实际上抓住了它的主要精神：信念被提高到如同均衡定义中策略那样的重要地位。形式上，新引进的"均衡"不再只是由每一个局中人的策略所组成，而且还包括了每个信息集上具有行动的每一个局中人的信念。以前我们坚持局中人应选择可信的策略，现在在

新的完美 Bayes 均衡中我们也坚持局中人拥有在均衡路径上与非均衡路径上的合理信念。非均衡路径上的信念问题将在稍后作出相应要求。

在下面两章我们将要介绍的信号博弈和廉价交谈这样的简单经济应用中，R_1—R_3 不仅抓住了完美 Bayes 精神而且构成了它的定义。但是，在较广泛一些的经济应用中，将需要更多的要求以删除难以置信的均衡。不同的作者使用不同的完美 Bayes 均衡定义。一个共同点是，所有定义均包含了 R_1、R_2、R_3 这三个要求，大多数定义还包括了对非均衡路径上信念的要求，这就是下面的 R_4；还有一些定义强加了更多的要求。本章遵循大多数定义的精神，在 R_1—R_3 要求之外再添加了如下 R_4：

R_4：在非均衡路径上信息集的信念通过 Bayes 法则和局中人的可能的均衡策略来确定。

关于"可能的均衡策略"的确切意思将在下面给予详尽阐述。我们先基于 R_1—R_4 定义完美 Bayes 均衡。

定义 12.1 *一个完美 Bayes 均衡由满足 R_1—R_4 的策略与信念构成。*

现在我们对 R_4 作一些详尽解释，所谓"可能的均衡策略"，这句话会使人产生概念上的模糊，在实际的经济应用中将针对具体情况加以说明，这里先观察一个三人博弈，见图 12.4。

这个三人博弈中，除了原博弈之外，还存在着一个真子博弈：它从局中人 2 的单结信息集开始，由于局中人 3 对局中人 2 的行动——L 还是 R——的信息不完美。因此相当于局中人 2 与局中人 3 之间的同时行动博弈，图 12.5 将这个真子博弈描绘成策略型。

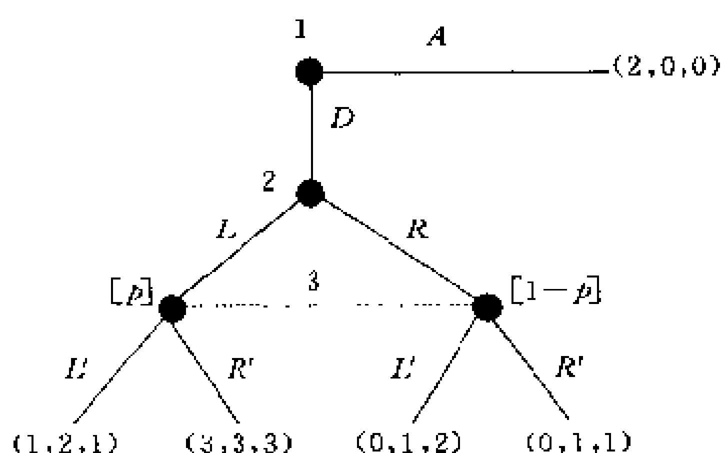


图12.4

局中人3

		L'	R'
局中人2	L	2,1	3,3
	R	1,2	1,2

图 12.5

图 12.5 中的盈利矩阵来自于图 12.4，其中各终点结处的是三元盈利向量中的元素，依次表示局中人 1、2、3 的盈利。显然，该真子博弈有唯一的 Nash 均衡(L,R')。因此，整个博弈无疑具有唯一的子博弈完美 Nash 均衡(D,L,R')。这个策略剖面是不是完美 Bayes 均衡呢？我们逐一验证定义 12.1 中的各项要求。根据 R_1，我们对局中人 3 的信息集赋予信念$(p,1-p)$（至于其他局中人行动时的信息集均为单结，信念是自然的）。从局中人 3 的观点看问题，倘若局中人 1 取行动 D，那么策略 R 是局中人 2 的劣策略，因而取 $p=1$，在给定如此的信念下，显然局中人 3 的最优选择是选取行动 R'。可见，(D,L,R')与信念 $p=1$ 满足要求 R_1—R_3。至于 R_4，我们完全不必担心，因为我们发现，在图 12.4 中不存在任何一个信息集不在该均衡途径上，这意味着要求 R_4“平凡地”得到满足。于是，(D,L,R')与信念 $p=1$ 构成图 12.4 中三人博弈的完美 Bayes

均衡。

图 12.4 中，对于子博弈完美 Nash 均衡(D,L,R')，R_4 的要求是“无所谓”的，因为不存在信息集在非均衡路径上，因此也不存在非均衡路径信息集上具行动的局中人的“可能均衡策略”。

现在我们再来考虑另一个策略剖面(A, L, L')以及信念 $p=0$。首先断定，这些策略是一个 Nash 均衡：没有一个局中人会单方面去偏离该策略剖面，因为信念 $p=0$，局中人 1 一旦偏离 A 将使他的盈利明显地减少。而当局中人 1 选定行动 A 后，事实上博弈已告结束，局中人 2 与 3 不存在什么偏离行为。其次，这些策略和信念也满足 R_1—R_3——每个局中人在其信息集上有信念。局中人 3 拥有信念 $p=0$，在给定这个信念下，局中人 3 的最优选择显然是 L'，这是策略剖面(A,L,L')中局中人 3 相应的策略行动。在预测到局中人 3 在后面将采取这个策略行动情况下，局中人 2 无疑取 L 才是最优选择，而在局中人 2 与 3 今后可能取(L,L')的条件下，局中人 1 的最优选择必定是 A。

Nash 均衡(A,L,L')肯定不是子博弈完美均衡，因为我们已经知道，博弈中仅有的子博弈拥有唯一的 Nash 均衡：(L,R')。既然(A,L,L')在真子博弈中的实现(L,L')不是该子博弈的 Nash 均衡，(A,L,L')当然不是子博弈完美 Nash 均衡。这个事实指出了，要求 R_1—R_3 不足以保证局中人的策略是子博弈完美 Nash 均衡。问题出在什么地方呢？读者不难发现，局中人 3 的信念 $p=0$ 与局中人 2 的策略 L 并不相合，因为 $p=0$ 蕴含着局中人 2 将取 R 而不是取 L。而我们“费尽心机”设立要求 R_1、R_2 与 R_3 并没有对局中人 3 的信念强加限制，如果博弈按照指定的策略(A,L,L')进行的话，局中人 3 的信息集不能达到。也就是说，局中人 3 的信息集不在 Nash 均衡(A,L,L')的路径上，关于该信息集上的信念 R_3 显得“无可奈何”。要求 R_4 可以用来处理这些非均衡路径上的信念确定问题。它使得我们可以使用局中人 2 的策略来确定局中人 3 的

信念，这里就出现了“局中人可能的均衡策略”的说法。这种说法虽然似乎模糊，但是它使得 R_4 可以对局中人 3 的信念强加限制：如果局中人 2 的策略是 L，那么局中人 3 的信念一定是 $p=1$；如果局中人 2 的策略是 R，那么局中人 3 的信念必定是 $p=0$。这里，非均衡路径信息集上的信念得到了确定，他们是通过我们的假设“如果……”来确定的，这个“如果……”就是描述了局中人的可能的均衡策略。通过 R_4 的“强行要求”，我们看到，倘若局中人 3 的信念相应于局中人 2 的可能均衡策略 L 是 $p=1$，那么 R_2 必定要求局中人 3 选取 R' 而不是 L'，这表明，策略 (A, L, L') 以及 $p=0$ 并不同时满足 R_1—R_4。因此，按照定义 12.1，(A, L, L') 和 $p=0$ 一起不能构成完美 Bayes 均衡。由于 R_4 的作用，我们排除了一个不合理的 Nash 均衡与信念：(A, L, L') 与 $p=0$，尽管它满足 R_1—R_3。

进一步阐述 R_4，假如图 12.4 博弈的情况发生了变化，在局中人 2 的单结信息集处“横生枝节”使局中人 2 有了第三个可能的行动 A'（见图 12.6，为讨论简单起见，在图中略去盈利函数）。

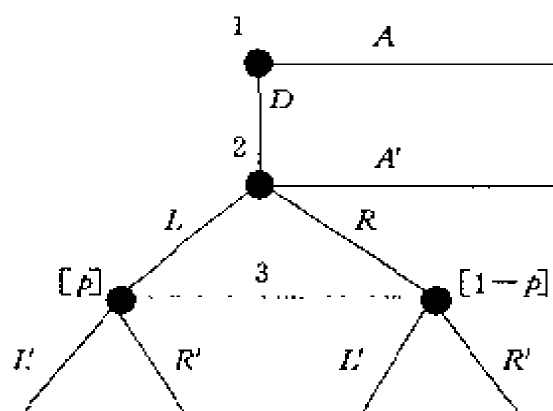


图 12.6

图 12.6 显示，若局中人 2 采取第三个可能行动 A'，那么博弈宣告结束。如前面所述那样，如果局中人 1 的均衡策略是 A，那么局中人 3 的信息集不在均衡路径上，博弈由图 12.4 变化为图 12.6 的最大区别之一在于，现在我们已经不能根据局中人 2 的可

能策略通过 R_4 去确定局中人 3 的信念了。如果局中人 2 的策略行动是 A'，显然 R_4 对局中人 3 的信念不可能产生任何限制，但是如果局中人 2 的策略是以概率 q_1 取 L，以概率 q_2 取 R、以概率 $(1-q_1-q_2)$ 取 A'，那么 R_4 指出局中人的信念为 $p=q_1/(q_1+q_2)$。这个变化了的博弈情况表明，在有些博弈场合，R_4 可能是无意义的。我们以后将要介绍的信号博弈就存在着 R_4 没有意义的现象，因为在那里，接收者根据发送者发出的信号树立已到达信息集上的信念，至于非均衡路径(即其他信号)的信念无法利用 R_4 来确定，从而在信号博弈中正式定义完美 Bayes 均衡时根本不考虑要求 R_4。我们简单地对"R_4 没有意义"作一描述：所谓 R_4 没有意义是指在非均衡路径上不能通过 R_4 对局中人的信念有所限制。

总结一下本节内容，在本节中我们引进了完美 Bayes 均衡概念，正如我们在本章开头所指出的，完美 Bayes 均衡实质上是对均衡的一种精炼。在 Nash 均衡里，每一个局中人的策略必定是关于其他局中人策略的最优反应，因此，没有一个局中人会选取严劣策略。而在完美 Bayes 均衡中，R_1 与 R_2 实际上等价于坚持没有一个局中人所采取的策略在任何信息集的开头是严劣的，信息集前的"任何"两字无疑表明这些信息集包含了均衡路径与非均衡路径上的信息集。我们知道，Nash 均衡与 Bayes Nash 均衡在非均衡路径的信息集上并不享有这个特征；如前面所指出的那样，即使子博弈完美 Nash 均衡在某些非均衡路径的信息集上也不享有这种特征，例如那些不包含在任何子博弈中的信息集。完美 Bayes 均衡堵塞了这些漏洞：局中人不能威胁取在非均衡路径的任何信息集起始时为严劣的策略。

完美 Bayes 均衡概念的长处之一是使得局中人的信念明确，并且因而允许我们不仅强加要求 R_3 与 R_4，而且对在非均衡路径上的信念(有时可以)强加进一步要求。由于完美 Bayes 均衡阻止局中人 i 取在非均衡路径信息集起始时为严劣的策略，因此局中

人 j 相信"局中人 i 将取严劣策略也许是不合理的"。但是，因为完美 Bayes 均衡使局中人的信念非常明确，因此，这样的均衡常常不能像以前我们构造子博弈完美均衡那样沿博弈树后退归纳而得到。R_2 部分地基于局中人在信息集上的信念而确定了局中人在该信息集的行动。如果 R_3 或者 R_4 的某一个应用于这个信息集，那么从博弈树较高处的局中人的行动确定了局中人的信念。可是 R_2 又部分地基于局中人以后的策略，反过来确定了博弈树较高处的行动。这种循环性意味着简单地通过博弈树后退工作通常不足以计算一个完美 Bayes 均衡。

§12.2 具可观察行动与不完全信息的多阶段博弈

完美 Bayes 均衡的最简单且具实际意义的例子当属信号博弈，由于信号博弈本身的重要性，我们将单独地另立一章进行详尽讨论。这里我们讨论更一般的博弈类，称之为"具可观察行动与不完全信息的多阶段博弈"。

由于考虑的是一般模型，因此我们假定博弈中存在 n 个局中人，每个局中人 i 的类型 θ_i 来自有限类型空间 Θ_i，令 $\theta=(\theta_1,\cdots,\theta_n)$。暂且假定类型互为独立，即当 $i\neq j$ 时，θ_i 与 θ_j 视作独立变量，因此关于类型向量 θ 的先验分布（通常认为由自然确定）p 实际上是各类型 θ_i 的边缘分布的乘积，即

$$P(\theta)=\prod_{i=1}^{n}p_i(\theta_i) \tag{12.3}$$

在博弈的一开始，每一个局中人知道自己的类型但是关于其对手的类型则没有任何信息。

我们已经接触过多阶段可观察行动博弈，它们分成若干阶段或周期进行，相应的博弈时间记作 $t=0,1,\cdots,T$，在每一周期 t，所有的局中人同时选取行动（或干脆不行动），且这些行动展示在该

周期的结尾以让下一个周期开始时为所有的局中人都能观察到。局中人决不接受关于 θ 的额外的观察结果。再假设每一个局中人在每一时刻的行动集独立于其类型。在定义此类博弈模型的完美 Bayes 均衡时，我们不得不再次启用以前多阶段博弈中的若干数学符号，例如，以 $a_i^t \in A_i(h^t)$ 表示局中人 i 在 t 时刻的行动，h^t 表示在 t 时刻行动前的历史。简记 $a^t=(a_1^t,\cdots,a_n^t)$ 以及 $h^t=(a^0,\cdots,a^{t-1})$。行为策略 σ_i 将可能的历史与类型集映射到行动空间：$\sigma_i(a_i|h^t,\theta_i)$ 表示给定 h^t 与 θ_i 时局中人 i 取行动 a_i 的概率。局中人 i 的盈利记为 $u_i(h^{T+1},\theta)$。

在完全信息可观察行动的多阶段博弈中，我们研究了子博弈完美均衡，要求策略剖面在每个真子博弈上都构成 Nash 均衡，在不完全信息可观察行动的多阶段博弈中，我们将这种思想进行推广，不仅要求策略剖面对于整体博弈形成完美 Bayes 均衡，而且对于每一个可能的历史 h^t 之后的每一个局期 t 开始的“后续博弈”也是如此。注意这一提法与完全信息时的“真子博弈”提法之间的差异，这正是完全信息动态博弈与不完全信息动态博弈的重要差异之一，完全信息动态博弈中的真子博弈开始于一个单结信息集，但是不完全信息中的后续博弈并非如此。因此，为了将后续博弈转换成真正的博弈，我们必须确定在每一个后续博弈开始时局中人的信念。仍沿用以前的记号，θ_{-i}表示局中人 i 的所有对手的类型向量，局中人 i 关于“认为其对手的类型为 θ_{-i}”的条件概率记为 $\mu_i(\theta_{-i}|\theta_i,h^t)$，$(i=1,2,\cdots,n;t=0,1,\cdots,T)$。

为定义完美 Bayes 均衡，R_1 是毋容置疑的，只要我们讨论对局中人 i 的信念应强制什么样的限制时，事实上已经确认了 R_1 的必要性。因此我们重点在于对信念的限制。注意到由自然确定局中人的类型分布为先验分布，类型为 θ_i 的局中人 i 在观察到历史 h^t 之后认定其他局中人的类型为 θ_{-i}的可能性显然是后验概率。类型互为独立的不完全信息博弈的经济应用通常作出如下假设，由

于假设是有关信念的，我们在各种条件（或限制）前加上字母 B。

$B(1)$ 后验信念为独立的，所有类型的局中人 i（即不管 θ_i 取 Θ_i 中何值）都有相同的信念，即，对所有 θ、t 及 h^t，

$$\mu_i(\theta_{-i}|\theta_i,h^t)=\prod_{j\neq i}\mu_i(\theta_j|h^t) \tag{12.4}$$

（12.4）式右边的条件（即"|"的后面部分）少了 θ_i 表示不管局中人 i 的类型如何，他认为局中人 $j(j\neq i)$ 的类型为 θ_j 的信念是相同的，（12.4）式右边这些后验边缘概率相乘则表示 $B(1)$ 所要求的后验信念的独立性。$B(1)$ 要求，即使未曾料到的观察结果也不能使局中人 i 相信他的对手的类型是相关的。

$B(2)$ 只要可能的话，使用 Bayes 法则修正 $\mu_i(\theta_j|h^t)$ 为 $\mu_i(\theta_j|h^{t+1})$：对所有 i，j，h^t，以及 $a_j^t\in A_j(h^t)$，如果存在满足下述条件的 $\hat{\theta}_j$：$\mu_i(\hat{\theta}_j|h^t)>0$，$\sigma_j(a_j^t|h^t,\hat{\theta}_j)>0$（后一条件表明在给定 h^t 时局中人 i 赋予 a_j^t 正概率），那么，对所有的 θ_j，

$$\mu_i(\theta_j|(h^t,a^t))=\frac{\mu_i(\theta_j|h^t)\sigma_j(a_j^t|h^t,\theta_j)}{\sum_{\tilde{\theta}_j}\mu_i(\tilde{\theta}_j|h^t)\sigma_j(a_j^t|h^t,\tilde{\theta}_j)} \tag{12.5}$$

注意：公式中左边条件 (h^t,a^t) 实际上就是 h^{t+1}。

（12.5）式形式上是 Bayes 公式，稍懂概率统计的读者对它都很熟悉。首先，$B(2)$ 指出，由于博弈在多阶段进行，每进行了一个周期，由于观察到的情况越多，供判断用的信息量也就越大，因此完全有可能利用 Bayes 法则对信念进行修正。（12.5）式指出，$B(2)$ 实际上比起以通常形式简单地使用 Bayes 法则要更强一些，因为它应用于当 t 周期时的历史 h^t 具有 0 概率时从 t 周期至 $(t+1)$ 周期的修正，也应用于当 h^t 具有正概率和某些局中人 $k\neq j$ 在 t 时刻选择 0 概率行动时关于局中人 j 的信念的修正。这两点从 $B(2)$ 的叙述立即可以看出，第一句"应用于"与叙述 $\mu_i(\hat{\theta}_j|h^t)>0$（只要存在这样的 $\hat{\theta}_j$）有关：h^t 的概率为 0，未必不存在 $\hat{\theta}_j$ 使得 $\mu_i(\hat{\theta}_j|h^t)>0$，诸如对非均衡路径信息集的讨论。而叙述"$\sigma_j(a_j^t|h^t,\hat{\theta}_j)>0$"

可以使我们立刻领会到在实际情况，当 $k \neq j$ 时该概率可以为 0，因为我们只需要存在 $\hat{\theta}_j$，不但 $\mu_i(\hat{\theta}_j|h^t)>0$，而且对于 j 有 $\sigma_j(a_j^t|h^t,\hat{\theta}_j)>0$ 就足够了。形成 $B(2)$ 的动机是，如果 $\mu_i(\cdot|h^t)$ 表示给定历史 h^t 时局中人 i 的信念，并且在 t 处不发生出人意外的事情，那么局中人 i 应当使用 Bayes 法则以形成在 $(t+1)$ 周期时它的信念。

$B(3)$　对所有 h^t,i,j,θ_j,a^t 与 $\hat{a}^t$，

$$\mu_i(\theta_j|(h^t,a^t))=\mu_i(\theta_j|(h^t,\hat{a}^t))\quad \text{如果 } a_j^t=\hat{a}_j^t \tag{12.6}$$

(12.6)式中，显然考虑的是情况 $a^t \neq \hat{a}^t$，$\mu_i(\theta_j|(h^t,\hat{a}^t))$ 表示了局中人 i 关于局中人 j 的信念从 t 周期到 $(t+1)$ 周期的修正结果，$B(3)$ 告诉我们，这种修正不受其他局中人行动的影响。

最后给出关于信念的一个限制，通常总是假定，当类型互为独立时，两个局中人 i 与 j 对于第三个局中人 k 的类型拥有相同的信念。

$B(4)$　对所有的 h^t,θ_k，以及 i，j，k 满足 $i \neq j, i \neq k, j \neq k$，

$$\mu_i(\theta_k|h^t)=\mu_j(\theta_k|h^t)=\mu(\theta_k|h^t) \tag{12.7}$$

$B(4)$ 蕴含了后验信念与给定 h^t 时 Θ 上的共同联合分布是一致的：

$$\mu(\theta_{-i}|h^t)\mu(\theta_i|h^t)=\mu(\theta|h^t)$$

读者不难发现，要求 $B(1)$—$B(4)$ 事实上对信息集上的信念确定作出了要求，从定义 Bayes 均衡的观点，它们相当于完成了 R_1、R_3、R_4 的任务。现在我们还需要 R_2 这样的要求，以保证那些满足 $B(1)$—$B(4)$ 的信念 μ 与策略 σ 在任意 t 与 h^t 以后的后续博弈中构成 Bayes 均衡。这就是我们下面提出的条件，我们用 (P) 来表示（以区别于 §1 中的 R_2）。

(P)　对于每一个局中人 i，类型 θ_i，局中人 i 有别于 σ_i 的另一个策略 σ_i'、以及历史 h^t，满足下述条件：

$$u_i(\sigma|h^t,\theta_i,\mu(\cdot|h^t))\geqslant u_i((\sigma_i',\sigma_{-i})|h^t,\theta_i,\mu(\cdot|h^t)) \tag{12.8}$$

现在我们可以定义具可观察行动与不完全信息的多阶段博弈中的 Bayes 均衡了。

定义 12.2 具可观察行动与不完全信息的多阶段博弈中的完美 Bayes 均衡包含了策略剖面 σ 与信念 μ，他们满足要求(P)与 $B(1)-B(4)$。

例 12.1 重复公共财产博弈

虽然我们已经正式地定义了具可观察行动与不完全信息的多阶段博弈中的完美 Bayes 均衡，但是举出一个具有普遍性的例子，的确使初学者会感到头痛。现只考虑简单一些的情况，以第十章中的公共财产提供问题作为阶段博弈重复进行二次。即 $t=0,1$，其中局中人 $i=1,2$。在每个 t 周期，局中人同时决定是否向该周期的公共财产作一点贡献，捐款是 0—1 决策的。在每个周期内，如果两人中至少有一人向公共财产捐款，则每个局中人获益 1，如果没有人捐款则获益为 0；局中人 i 在两个周期中的捐款费用均为 c_i。每周期的盈利情况重新描述如图 12.7。

		局中人 2 捐款	局中人 2 不捐款
局中人 1	捐款	$1-c_1, 1-c_2$	$1-c_1, 1$
	不捐款	$1, 1-c_2$	0,0

图 12.7

假定盈利考虑周期的折扣因素，因此两次博弈的总盈利应当为第一周期盈利($t=0$ 时)加上第二周期($t=1$)盈利的 δ 倍($0<\delta<1$)。从公共财产处每次受益为 1(只要有人捐款)是博弈的共同知识，但是每个局中人的捐款费用只有他自己知道。然而，两个局中人都相信 c_i 独立地来自 $[0,\bar{c}]$($\bar{c}>1$)上连续且严增的累积分布函数 $p(\cdot)$。

从第十章中我们已经知道，如果博弈只进行一次，并且如果方

程 $c^*=1-p(1-p(c^*))$有唯一解，那么博弈有唯一 Bayes 均衡，且 $c^*=1-p(c^*)$。类型 $c_i\leqslant c^*$ 则捐款而其余类型不捐款。

这里我们将考虑再重复博弈一次，每个周期每个局中人的行动集为{0,1}。局中人 $i(i=1,2)$的策略由一对概率 σ 组成：$\sigma_i^0(1|c_i)$——局中人 i 的成本费用为 c_i 时他在第一次博弈中捐款的概率，和 $\sigma_i^1(1|h^1,c_i)$ ——在第二次博弈时，当他的成本为 c_i 且历史为 h^1 时局中人 i 捐款的概率，其中 h^1 有四种可能：{00，01，10，11}。该集合中每个元素由两个 0 或 1 的数字组成，第 1 个数字表示局中人 1 在第一次博弈时的决策，第 2 个数字则表示局中人 2 在第二次博弈时的决策。

现令 z_j 表示局中人 j 第一次捐款的概率，而以 z_j^{xy} 表示在第一次两人的捐款决策的基础上，局中人 j 在第二次捐款的条件概率，所谓条件 xy 是指：x 表示局中人 i 在第一次是否捐款(x 取 0 或 1)，y 则表示局中人 j 在第一次是否捐款(y 取 0 或 1)。以局中人 1 为例，如果局中人 1 在第一次不捐款，他在两次博弈中获得的期望盈利为

$$z_2+\delta\{(1-c_1)(z_1^{00}+z_1^{01})+(z_2^{00}+z_2^{01})\} \tag{12.9}$$

如果局中人 1 在第一次捐款，他在两次博弈中获得的期望盈利为

$$(1-c_1)+\delta\{(1-c_1)(z_1^{10}+z_1^{11})+(z_2^{10}+z_2^{11})\} \tag{12.10}$$

当且仅当(12.10)式大于等于(12.9)式时，局中人 1 在第一次博弈时捐款，即

$$(1-c_1)+\delta(1-c_1)(z_1^{10}+z_1^{11}-z_1^{00}-z_1^{01})\geqslant z_2+\delta(z_2^{00}+z_2^{01}-z_2^{10}-z_2^{11}) \tag{12.11}$$

故

$$c_1\{1+\delta(z_1^{10}+z_1^{11}-z_1^{00}-z_1^{01})\}$$
$$<\{1+\delta(z_1^{10}+z_1^{11}-z_1^{00}-z_1^{01})\}-\{z_2+\delta(z_2^{00}+z_2^{01}-z_2^{10}-z_2^{11})\} \tag{12.12}$$

事实上，由(12.12)式，我们基本上证明了对每一个局中人 i，必定

存在一个 $\hat{c}_i$，使得，当且仅当 $c_i \leqslant \hat{c}_i$ 时局中人 i 在第一周期捐款（至少在某些特殊情况可以做到，例如，δ 比较小等）。其实我们还可以证明 $\hat{c}_i$ 满足 $0<\hat{c}_i<1$。

我们现在只考虑寻求对称的完美 Bayes 均衡，即 $\hat{c}_1=\hat{c}_2=\hat{c}$。通过在给定的后验信念下解第二个周期的 Bayes 均衡开始这个过程。该后验信念由均衡策略与第一个周期的结局得到确定。

(1)没有一个局中人在第一个周期捐款

在这样的第一周期结局下，两个局中人获悉各自的对手的成本超过 $\hat{c}$。后验累积信念于是为

$$P(c_i|00)=\frac{P(c_i)-P(\hat{c})}{1-P(\hat{c})} \quad c_i\in[\hat{c},\bar{c}] \tag{12.13}$$

$$P(c_i|00)=0 \quad c_i\leqslant\hat{c} \tag{12.14}$$

(12.14)式是显然的，因为在“00”下，不可能小于 $\hat{c}$，否则将发生局中人 i 捐款现象。至于 $c_i\in[\hat{c},\bar{c}]$，只要考虑图 12.8。



图12.8

“00”表示仅需考虑区间$[\hat{c},\bar{c}]$，条件累积信念显然为线段$[\hat{c},c_i]$的长度与线段$[\hat{c},\bar{c}]$的长度之比，故得(12.13)式。

在对称的第二周期均衡中，如第十章所讨论的那样，在$[\hat{c},\bar{c}]$中必存在一个临界点 c^0；当且仅当 $\hat{c}\leqslant c_i\leqslant c^0$ 时局中人 i 在第二周期捐款，根据那里的讨论，c^0 应等于对手不捐款的概率，注意这里所谓概率当然是指条件概率，即在$[\hat{c},\bar{c}]$上进行计算，易知

$$c^0=(1-P(c^0))/(1-P(\hat{c})) \tag{12.15}$$

显然 c^0 满足 $\hat{c}<c^0<1$。如果在第一周期没有一个人捐款，类型 $\hat{c}$ 的局中人在第二周期将会捐款，且他在第二周期的盈利必然等于 $v^{00}(\hat{c})=1-\hat{c}$。读者应当注意这样的事实，尽管局中人 i 在两个周期类型相同，即捐款成本 c_i 相同，但在两个周期内的中间点或临

界点并不相同。

(2)两个局中人在第一周期都捐款

此时后验累积分布为

$$P(c_i|11)=\frac{P(c_i)}{P(\hat{c})} \quad c_i\in[0,\hat{c}] \tag{12.16}$$

$$P(c_i|11)=0 \quad c_i\in[\hat{c},\bar{c}] \tag{12.17}$$

这是因为"11"条件蕴含着两个局中人都知道对手的成本 c_i 不超过 $\hat{c}$ 的缘故。在对称的第二周期均衡中，每个局中人 i 当且仅当 $c_i\leqslant\tilde{c}$时捐款，$\tilde{c}$ 满足 $0<\tilde{c}<\hat{c}$。与(1)中讨论的同样理由，每个局中人的中间点成本等于他的对手不捐款的条件概率，即

$$\tilde{c}=\frac{P(\hat{c})-P(\tilde{c})}{P(\hat{c})} \tag{12.18}$$

注意到一个事实，类型 $\hat{c}$ 此时不捐款，因此他的第二周期盈利是 $v^{11}(\hat{c})=P(\tilde{c})/P(\hat{c})$。读者仍会发现，第二周期的中间点 $\tilde{c}$ 不同于第一周期的 $\hat{c}$。

(3)只有一个局中人在第一周期捐款

假设局中人 i 在第一个周期捐款而此时局中人 $j(\neq i)$不捐款，这表明 $c_i\leqslant\hat{c}\leqslant c_j$。那么在第二个周期内的一个均衡是：局中人 i 捐款及局中人 j 不捐款，此时类型 $\hat{c}$ 的第二周期盈利分别为 $v^{10}(\hat{c})=1-\hat{c}$ 与 $v^{01}(\hat{c})=1$。

现在需要考虑第一个周期的均衡了，$\hat{c}$ 作为捐款或不捐款的临界点，必然会有类型为 $\hat{c}$ 的局中人，对于选择捐款还是选择不捐款应表现出"无所谓"态度，或者说，这两种选择为他带来一样多的盈利：

$$\begin{aligned}&1-\hat{c}+\delta\{P(\hat{c})v^{11}(\hat{c})+[1-P(\hat{c})]v^{10}(\hat{c})\}\\&=P(\hat{c})+\delta\{P(\hat{c})v^{01}(\hat{c})+[1-P(\hat{c})]v^{00}(\hat{c})\}\end{aligned} \tag{12.19}$$

将上述讨论的三种情况中的盈利函数 $v^{00}(\hat{c})$、$v^{11}(\hat{c})$、$v^{01}(\hat{c})$、$v^{10}(\hat{c})$分别代入(12.19)式，并注意到(12.18)式，经整理得

$$1-P(\hat{c})=\hat{c}+\delta P(\hat{c})\tilde{c} \tag{12.20}$$

回想如果公共财产问题仅博弈一次，捐款临界点 c^* 满足 $c^*=1-P(c^*)$，为了能与 $\hat{c}$ 作比较，利用(12.20)式可得

$$\begin{cases} c^*+P(c^*)=1 \\ \hat{c}+P(\hat{c})(1-\delta\tilde{c})=1 \end{cases} \tag{12.21}$$

由于 $\tilde{c}>0,\delta>0$，又 $P(c)$是 c 的严格递增函数，因此必有

$$\hat{c}<c^* \tag{12.22}$$

这表明如果类型 $c\in(\hat{c},c^*)$，那么该类型 c 在一次性公共财产博弈的均衡结局中应捐款，而在二次博弈的第一个周期中不必捐款。

§12.3 序贯均衡

本节将讨论 Kreps 与 Wilson 的序贯均衡(sequential equilibrium)，这个概念是为"一般展开型博弈"而定义，它比起完美 Bayes 均衡来对零概率事件上的信念置于更多的限制。在序贯均衡中，局中人的信念为仿佛在每一个信息集上存在着小概率的"颤抖"或"失误"。我们将看到，除非博弈至多有两个周期或者每个局中人至多有两个类型，一般地序贯均衡比起完美 Bayes 均衡来更强一些。

为讨论新的概念，我们先考虑具有有限(n)个局中人和有限个决策结的完美回忆博弈，并规定若干记号以便于问题的展开。其实这些记号与刚学习展开型博弈时并无太多不同，只不过它们都以决策结 x 作为出发点。例如，包含结 x 的信息集以 $h(x)$记之；在结 x 处行动的局中人 i 则记作 $i(x)$；终点结仍以 z 表示。局中人 $i(x)$ 在结 x 处的混合或行为策略记作 $\sigma_i(\cdot|x)$或 $\sigma_i(\cdot|h(x))$。$\sum$表示所有策略剖面 $\sigma=(\sigma_1,\cdots,\sigma_n)$的集合全体，"自然"行动的外生概率分布(先验分布)以 p 表示。一个重要的事实是，迄今为止我们

所考虑的自然行动不是由策略给出的，因此，当我们以“颤抖”扰动博弈时，自然行动本身不会受到影响。

当策略剖面 σ 为给定时以 $P^{\sigma}(x)$ 与 $P^{\sigma}(h)$ 分别表示结 x 和信息集 h 达到的概率。这些概率显然依赖于先验分布 p，只不过由于每一个给定的展开型博弈中先验概率 p 是确定了的。因此 $P^{\sigma}(x)$ 与 $P^{\sigma}(h)$ 中不再标以 p。信念体系 μ 确定了在每个信息集 h 上的信念：$\mu(x)$ 表示局中人 $i(x)$，基于信息集 $h(x)$ 已到达所赋予结 x 的条件概率。为说明 $P^{\sigma}(\cdot)$ 与 $\mu(x)$ 等概念，以博弈树（见图 12.9）为例。

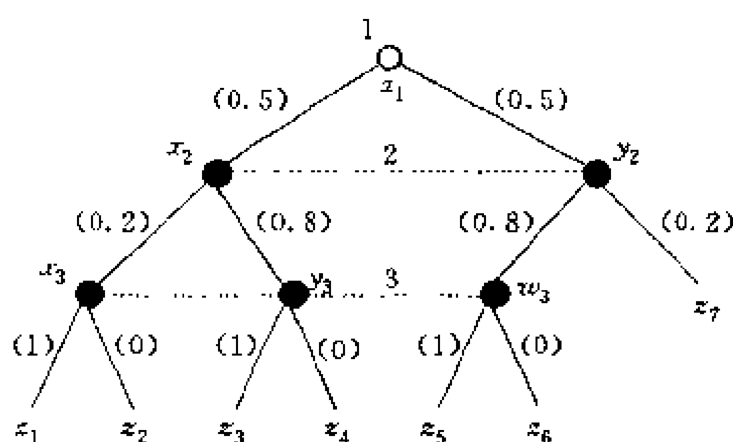


图12.9

我们计算相应的 $P^{\sigma}(x)$、$P^{\sigma}(h)$ 与 $\mu(x)$ 如下：$P^{\sigma}(x_1)=1$ 是显然的；$P^{\sigma}(x_2)=P^{\sigma}(y_2)=0.5$ 如图 12.9 所示。需要稍作解释的是以下计算内容：

$$P^{\sigma}(x_3)=0.5\times0.2=0.1$$

$$P^{\sigma}(y_3)=P^{\sigma}(w_3)=0.5\times0.8=0.4$$

$$P^{\sigma}(h\ )=P^{\sigma}(y_3)+P^{\sigma}(x_3)+P^{\sigma}(w_3)=0.4+0.1+0.4=0.9$$

$$h=\{x_3,y_3,w_3\}$$

$$P^{\sigma}(z_1)=P^{\sigma}(z_7)=0.5\times0.2=0.1$$

$P^{\sigma}(z_2)=P^{\sigma}(z_4)=P^{\sigma}(z_6)=0$ 为显然

$$P^{\sigma}(z_3)=P^{\sigma}(z_5)=0.5\times0.8\times1=0.4$$

$\mu(x_3)=P^\sigma(x_3)/P^\sigma(h)=0.1/0.9=1/9$

$\mu(y_3)=\mu(w_3)=0.4/0.9=4/9$

博弈的盈利当然由终点结处确定，如果终点结 z 到达，局中人 i 的盈利仍记作 $u_i(z)$。现令 $u_{i(h)}(\sigma|h,\mu(h))$表示局中人 $i(h)$(即在信息集 h 上有行动的局中人 i)在信息集 h 到达，局中人的信念为 $\mu(h)$以及策略为 σ 的情况下的期望盈利(或期望效用)。

在前面讨论的完美 Bayes 均衡时我们提到不但要有策略而且应有信念，将这种思路推广到一般展开型博弈时，这两个要素自然不可缺少。我们称(σ,μ)为一个状态(assessment)，它确定了一个策略剖面 σ 与一个信念体系 μ，所有可能状态的集合记作 Ψ。

用图 12.9 来阐述上面引进的记号之外，还有一个目的是，我们看到这表面上看似较复杂的博弈树，除了整体博弈之外不存在其他的真子博弈，因而要求局中人的策略在每一个子博弈中形成 Nash 均衡似乎在某些情况成为一句空话，或者说这种要求太弱了一些。实际情况是，在不完全信息或不完美信息的博弈中只有很少真子博弈，因此我们有必要对前面所讨论的完美 Bayes 均衡中的要求作一些改进或者精炼。

信念体系实际上是对 R_1 的重申，我们的讨论总是在给定信念体系下进行的，因此有时不另外特地去提 R_1。至于 R_2，要求在给定信念下，局中人的策略是序贯理性的，用本节引进的记号与说法就是，给定了信念体系，没有一个局中人可能通过任意信息集上的偏离而获益。由于要求策略与信念是序贯理性(sequentially rational)，因而我们将该条件索性记作(S)：

(S)　状态(σ,μ)如果满足，对任意信息集 h 与另外的策略 $\sigma'_{i(h)}$，使得

$$u_{i(h)}(\sigma|h,\mu(h))\geqslant u_{i(h)}((\sigma'_{i(h)},\sigma_{-i(h)})|h,\mu(h)) \quad (12.23)$$

那么称(μ,σ)是序贯理性的。

注意到序贯理性条件(S)是针对状态(μ,σ)的，它由所有局中

人的策略剖面和所有信息集上的概率分布组成。因此，(S)可以视作多阶段博弈中条件(P)的适当推广，如果我们考虑的不是一般的展开型博弈、而是多阶段博弈，读者不难发现，此时(S)实际上就是(P)。接下来我们应当对信念提出要求，如前所述知，最困难也是最易发生争议的地方就是对非均衡路径信息集上信念的限制条件，对完美 Bayes 均衡的精炼，许多功夫化在非均衡路径信息集上。我们顺着 Kreps 与 Wilson 的办法首先定义一致性(consistency)：

令$\sum^0$表示所有完全混合(行为)策略的集合，即，$\sum^0$中的策略剖面σ满足对所有h与$a_i \in A(h)$成立$\sigma_i(a_i|h)>0$。若$\sigma\in\sum^0$，那么对所有x结有$P^\sigma(x)>0$，因而$\mu(x)=P^\sigma(x)/P^\sigma(h(x))$，关于图 12.9 中的$\mu(x_3)$、$\mu(y_3)$与$\mu(w_3)$，就是这样计算而得的。令$\Psi^0$表示所有满足下述条件的状态$(\sigma,\mu)$的集合：$\sigma\in\sum^0$，$\mu$(唯一地)通过 Bayes 法则由$\sigma$所确定。一致性条件用其英文名称的第一个字母($c$)表示：

(c) 状态(σ,μ)称为一致性的，如果存在Ψ^0中的某状态序列(σ^m,μ^m)，使得

$$(\sigma,\mu)=\lim_{m\to\infty}(\sigma^m,\mu^m) \tag{12.24}$$

这里，策略σ不必完全是混合的，但是，他们及其相应的信念可以认为是完全混合的策略与有关信念的极限。熟悉极限理论的读者十分清楚，若数列y_m的极限是y_0，y_m在y_0的左右(或上下)波动并随着m的增大波幅越来越小地接近y_0。形象一点地讲，那些y_m可以视作(或理解为)y_0产生的"颤抖"。状态(σ,μ)的一致性，也可以将Ψ^0中的状态序列(σ^m,μ^m)理解为(σ,μ)的"颤抖"。注意到一致性定义涉及到策略，自然行动上的概率分布(先验分布)并不是用策略来表示的，一致性定义不能将"颤抖"应用于自然行动。

有了序贯理性条件(S)和一致性条件(c)，我们可以定义序贯均衡。

定义 12.3 满足条件(S)与(c)的状态(σ,μ)称为序贯均衡。

将(P)推广到序贯理性条件(S)是容易使人理解的，因为一般展开型博弈与多阶段博弈之间毕竟存在着差异。但是在对信念需要加限制时，Kreps 与 Wilson 怎么会想到引进“一致性”这个概念呢？我们对他们的动机很感兴趣。先来观察他们给出的一个简单的博弈示意图（见图 12.10）。

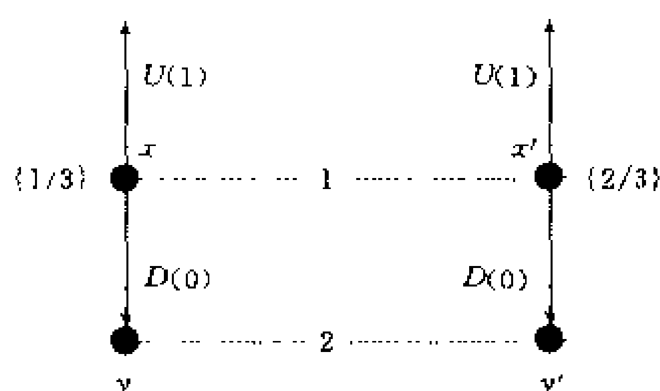


图12.10

如图 12.10，局中人 1 赋予结 x 以1/3概率，赋予结 $x'\in h(x)$ 以2/3概率，即 $\mu(x)=1/3,\mu(x')=2/3$，不管在 x 还是在 x'，他的策略是选择 U。因为在结 x 或 x' 处，图示取 U 的概率为 1 而 D 的概率为 0。显然局中人 2 的信息集（包括结 y 与 y'）为非均衡路径。如果局中人 1 偏离均衡而采取了行动 D，那么局中人 2 的信念又如何呢？由于局中人 1 不能区分 x 与 x'，因而很自然地要求他在这两个结上都有“可能”偏离。于是局中人 2 将置1/3概率于 y 而置2/3概率于 y'，或 $\mu(y)=1/3,\mu(y')=2/3$。但是，由于图 12.10 中采取 D 实际上在均衡中是一个零概率事件，因而我们取随便怎样的 $\mu(y)$ 都与 Bayes 法则保持一致。换句话说，局中人 2 在非均衡路径信息集上拥有任何信念 $\mu(y)$ 都是可以的。注意到零概率事件

并不等于“绝对不可能发生”的事件，它是有可能发生的。在这种情况下，认为“局中人 2 的信念 $\mu(y)$随便等于什么都可以”未免欠妥当。Kreps 与 Wilson 引进的“一致性”则产生了该博弈的“正确信念”。令 $0<\epsilon<1$，则 ϵ^m 为收敛于 0 的正数序列（m 取自然数），我们将 ϵ^m 解释为局中人 1“颤抖”且取 D 的概率。对于这样一个数列

$$\mu^m(y)=\frac{\mu(x)\epsilon^m}{\mu(x)\epsilon^m+\mu(x')\epsilon^m}=\frac{1}{3} \tag{12.25}$$

对(12.25)式稍作解释：由于 ϵ^m 视作局中人 1 颤抖且取行动 D 的概率系列，因此从 x 结“颤抖”地到达 y 的概率为 $\mu(x)\epsilon^m$，从 x' 结“颤抖”地到达 y' 的概率为 $\mu(x')\epsilon^m$，局中人 1 不能区分 x 与 x'，因此局中人 2 相信“颤抖”地到达 y 的概率应为形如(12.25)式的条件概率。于是，颤抖保证了局中人的信念考虑到了信息结构。通过颤抖，我们发现原先所讲的局中人 2 的信念 $\mu(y)=1/3$ 是合乎情理的，故也认为它是正确的，我们不必再考虑任何其他的信念 $\mu(y)$，尽管它们与 Bayes 法则保持着一致。

对序贯均衡作一个小结，序贯均衡是建立在条件(S)与(c)上的状态(σ,μ)，既然序贯理性条件(S)是（多阶段博弈中的）条件(P)的适当推广，而一致性条件(c)又能通过“颤抖”妥善处理非均衡路径信息集上的信念，因此一般地说来，序贯均衡是比完美 Bayes 均衡更强的概念。基于此，我们需要略微了解一下序贯均衡的性质以及它与完美 Bayes 均衡之间的比较。

(1)存在性

像 Nash 均衡的存在性一样，我们要问序贯均衡在展开型博弈中是否存在。这里，我们不加证明地指出，在任何有限展开型博弈中，至少存在一个序贯均衡。

(2)上半连续性

序贯均衡对应关于盈利（或效用）是上半连续的，有如下结论：

对于任何一个收敛于某 u 的盈利函数序列 u^m(u^m、u 所相应的

博弈仍记为 u^m 与 u)，倘若对所有 m，状态(σ^m,μ^m)是博弈 u^m 的序贯均衡，并且(σ^m,μ^m)收敛于(σ,μ)，那么状态(σ,μ)是博弈 u 的序贯均衡。

(3)序贯均衡的结构

定理 12.1 (Kreps & Wilson 1982)对于普通的完美回忆有限展开型博弈(所谓"普通的"是指具有通常的终点盈利)，终点结上的序贯均衡规律分布集合是有限的。

即对于固定的展开型与固定的先验信念，使得有关博弈具有无限个序贯均衡结局的盈利函数 u 的集合，其闭包的 Lebesgue 测度为零。由于在确定非均衡路径信念上存在一定的余地，一般地，序贯均衡状态集是个无限集。相应地，序贯均衡策略集合一般地也是无限的，因为，当局中人在非均衡路径信息集的两个行动之间感到无所谓时，可以确定该信息集上许多不同的随机概率。因此，根据定理 12.1，我们可以知道，尽管序贯均衡结局的集合是有限的，正因为在非均衡路径信息集上信念确定上的"富有余地"，一般地，序贯均衡状态却是无限多个。

(4)序贯均衡与完美 Bayes 均衡的比较

诚如序贯均衡概念引出时所说，在展开型博弈中，序贯均衡定义"强"于完美 Bayes 均衡，或者说，序贯均衡是完美 Bayes 均衡的精炼。但是，验证一个状态(σ,μ)为序贯均衡是件困难的事情。尤其是对一致性条件(c)的验证，人们必须找到 Ψ^0 中的(σ^m,μ^m)序列。举一个简单的例子(见图 12.11)。

在图 12.11 所展示的博弈树中，可能的序贯均衡是 A 与(R_1，R_2)。我们仅考虑局中人 1 取 A 的结局。将验证状态($\sigma_1(A)=1$，$\sigma_2(L_2)=1$；$\mu(w)=1$)是序贯均衡。也就是要验证它满足条件(S)与(c)。状态(S)满足是显然的，因为在 $\sigma_2(L_2)=1$ 情况下，局中人 1 不可能偏离行动 A、否则他将至多获盈利 1。而当局中人 1 取定行

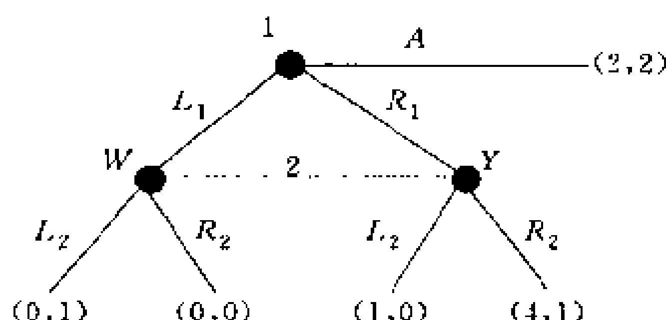


图12.11

动 A 时，博弈已告结束。关键在于一致性条件(c)是否满足。考虑如下颤抖：

$$\sigma_1^m(A)=1-\frac{1}{m}-\frac{1}{m^2}$$

$$\sigma_1^m(L_1)=\frac{1}{m},\sigma_1^m(R_1)=\frac{1}{m^2}$$

$$\sigma_2^m(L_2)=1-\frac{1}{m},\sigma_2^m(R_2)=\frac{1}{m}$$

显然 $\sigma^m=(\sigma_1^m,\sigma_2^m)$是完全混合策略。即 $\sigma^m\in\sum^0$，由 Bayes 法则，$\mu^m(w)=\frac{1}{m}/(\frac{1}{m}+\frac{1}{m^2})=\frac{m}{m+1}\to 1$ （当 $m\to\infty$)时。因此$(\sigma^m,\mu^m)\to(\sigma,\mu)$ （当 $m\to\infty$)时。

这是一个非常简单的例子，因此在构造(σ^m,μ^m)时比较容易，即便这样，对于一般的财经类读者在此类“构造性”技巧中将感到困难，更别说稍微复杂一些的经济学方面的应用了。

针对这个例子我们顺便提到一个有趣的事实：当博弈添加一个显而易见不相干的行动或策略，序贯均衡集合可能发生改变。例如，图 12.11 若变成图 12.12，与图 12.11 相比，图 12.12 添加了一个不相干行动 NA，其余的一切不变。但是由于 NA 的添加，这里出现了除原博弈之外的“同时行动”真子博弈，它发生在 NA 之后。这个子博弈的唯一 Nash 均衡是(R_1,R_2)，因为从局中人 1 的

角度出发，R_1 严优于 L_1。因而，博弈的唯一序贯均衡盈利向量应为(4,1)而不是(2,2)，局中人 1 不会舍弃盈利 4 而去取盈利为 2 的行动 A。因此，A 不是序贯均衡结局。

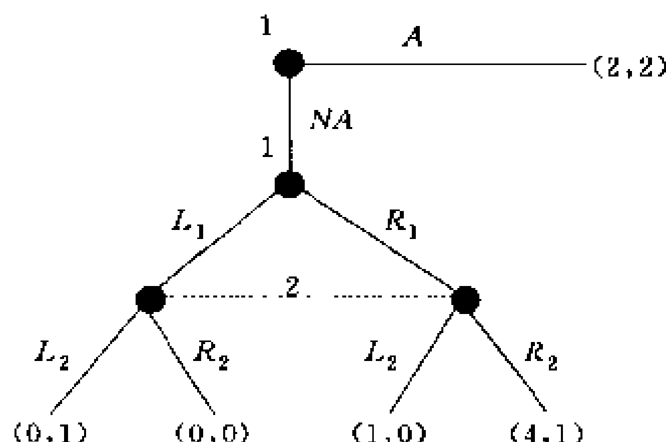


图12.12

现在回到序贯均衡与完美 Bayes 均衡的比较，既然序贯均衡的验证有一定难度，如果在某些场合，序贯均衡等价于完美 Bayes 均衡的话，那么人们宁可使用完美 Bayes 均衡概念，这种现象在许多经济应用中存在。1991 年 Fudenberg 与 Tirole 对不完全信息的多阶段博弈证明了一个这样的事实：

定理 12.2 (Fudenberg & Tirole 1991)考虑具有独立类型的不完全信息多阶段博弈。如果每个局中人至多有两个可能类型，或者博弈有两个周期，多阶段博弈中的条件(B)等价于一致性(C)，因此，完美 Bayes 均衡与序贯均衡这两个集合为一致。

倘若每个局中人的类型多于两个，或者博弈的周期数不止两个，条件(B)不再足以得到一致性。图 12.13 阐述了这个可能性。图 12.13 中，局中人 1 有三种类型：θ_1'、θ_1''与 θ_1^*，但是在 t 时刻，来自以前行动的 Bayes 推断导致局中人 1 必定属类型 θ_1^* 的结论。在该点处的均衡策略在图中用圆括号标出，即类型 θ_1'取行动 a_1'，θ_1''类型取行动 a_1''，以及类型 θ_1^* 取行动 a_1^*（图中类型 θ_1^* 的零概率行

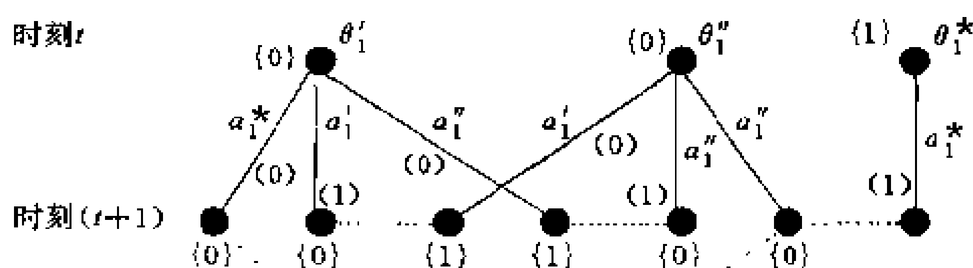


图12.13

动没有标绘出来)。由于θ_1'、θ_1''具有0概率,故局中人2预期看到局中人1采取行动a_1^*。现在我们关心的是非均衡路径上局中人2的信念,就是指如果局中人2看到了a_1'与a_1''中的一个出现,此时他的信念如何?该信念在图中用花括号标出,即,如果局中人2看到a_1',他认为局中人1的类型为θ_1'',而a_1''是局中人1为类型θ_1'的信号。由于完美Bayes均衡的定义对刚发生偏离的局中人的信念没有约束,因此图12.13中的情况与完美Bayes均衡相一致。即图12.13所发生的情况与完美Bayes均衡没有矛盾与冲突。

然而,这种情况不可能成为序贯均衡的一部分。为看到这一点,设想存在着颤抖σ^m收敛于给定的策略σ,以及与颤抖相联的信念μ^m收敛于给定的信念μ,令μ^m在时刻t赋予类型θ_1'的概率为$(\epsilon')^m$,赋予类型θ_1''的概率为$(\epsilon'')^m$。因为μ^m收敛于μ、$(\epsilon')^m$与$(\epsilon'')^m$都收敛于0,且$\sigma^m(a_1'|\theta_1'')$与$\sigma^m(a_1''|\theta_1')$也收敛于0。因为

$$\mu^m(\theta_1''|a_1')=\frac{\mu^m(\theta_1'')\sigma^m(a_1'|\theta_1'')}{\sum_{\theta_1}\mu^m(\theta_1)\sigma^m(a_1'|\theta_1)} \tag{12.26}$$

因此

$$\mu^m(\theta_1''|a_1')=\frac{(\epsilon'')^m\sigma^m(a_1'|\theta_1'')}{(\epsilon')^m\sigma^m(a_1'|\theta_1')+(\epsilon'')^m\sigma^m(a_1'|\theta_1'')}$$
$$=1\Big/\left\{\frac{(\epsilon')^m}{(\epsilon'')^m}\cdot\frac{\sigma^m(a_1'|\theta_1')}{\sigma^m(a_1'|\theta_1'')}+1\right\} \tag{12.27}$$

由于$\sigma^m(a_1'|\theta_1'')$收敛于0,且$\sigma^m(a_1'|\theta_1')$收敛于1,因而若要$\mu^m(\theta_1''|a_1')$收敛于1(即图中所示花括号的后验信念),则必须$(\epsilon')^m/$

$(\epsilon'')^m$ 收敛于 0。这表明，如图 12.13 所示，当类型 θ_1' 以概率 1 取行动 a_1' 而类型 θ_1'' 以概率 0 取行动 a_1' 时，为了要使行动 a_1' 之后的信念"集中"在类型 θ_1'' 上（即 $\mu^m(\theta_1''|a_1')\to 1$），必须 $(\epsilon')^m$ 比起 $(\epsilon'')^m$ 来是更高阶的无穷小量，用具体情况来描述，就是在 t 时刻的"先验"信念必定是类型 θ_1'' 比起类型 θ_1' 有无穷多可能。单独地看问题，这个要求与序贯均衡相一致。但是，如果我们同时考虑 $\mu^m(\theta_1'|a_1'')$ 收敛于 1 的问题，同样的讨论导致要求 $(\epsilon'')^m/(\epsilon')^m$ 收敛于 0，就是说，要求 t 时刻的"先验"信念必定是类型 θ_1' 比起类型 θ_1'' 来有无穷多可能，这两个条件放在一起与需要保持一致的信念是不相容的。可见，图 12.13 的状况不可能是序贯均衡的一部分。

上面所述的例子提示我们，局中人的类型不止两个时，存在着零概率类型不止一个的情况，零概率类型的相对概率（即上例中的 $(\epsilon')^m/(\epsilon'')^m$ 与 $(\epsilon'')^m/(\epsilon')^m$）看来应有一定的限制才能与一致性条件相容。因此，为了使完美 Bayes 均衡具有一致性，必须延拓信念的定义以控制零概率类型的相对概率，并且必须对这些相对概率调整的方式强加一定的限制。要求局中人评估自然的零概率状态相对概率是过分了些，但是容易使其公式化。形式地，人们希望关于每个局中人的后验信念形成"相对信念体系"或"条件概率体系"。即，纵然在历史 h^t 条件下，局中人 i 可能为类型 θ_i 或 θ_i' 都是零概率事件，局中人关于"在历史 h^t 与局中人 i 可能为类型 θ_i 或 θ_i' 条件下，局中人 i 具有类型 θ_i"事件具有信念 $\mu^*(\theta_i|(\theta_i,\theta_i'),h^t)$。注意到，相对信念体系通过公式 $\mu(\theta_i|h^t)\equiv\mu^*(\theta_i|\Theta_i,h^t)$ 产生绝对信念体系。我们称 (σ,μ^*) 为广义状态（generalized assessment）。

现在我们的任务是，将根据不完全信息与可观察行动的多阶段博弈里完美 Bayes 均衡中的 Bayes 条件 $B(1)-B(4)$ 延拓以要求 Bayes 法则和无信号条件不仅对绝对信念成立，而且对相对信念也成立。

(B^*)　一个广义状态(σ,μ^*)满足条件(B^*),如果

$B^*(1)$　只要有可能,用 Bayes 法则修正信念 $\mu^*(\theta_i|(\theta_i,\theta_i'),h^t)$ 为 $\mu^*(\theta_i|(\theta_i,\theta_i'),(h^t,a^t))$,即,如果 a_i^t 在条件(θ_i,θ_i')与历史 h^t 下具有正概率,那么

$$\mu^*(\theta_i|(\theta_i,\theta_i'),(h^t,a^t))=\frac{\mu^*(\theta_i|(\theta_i,\theta_i'),h^t)\sigma_i(a_i^t|h^t,\theta_i)}{\sum_{\tilde{\theta}_i=\theta_i,\theta_i'}\mu^*(\tilde{\theta}_i|\theta_i,\theta_i'),h^t)\sigma_i(a_i^t|h^t,\tilde{\theta}_i)}\tag{12.28}$$

$B^*(2)$　后验信念是独立的:

$$\mu(\theta|h^t)=\prod_i\mu(\theta_i|h^t)\tag{12.29}$$

$B^*(3)$　关于局中人 i 在时刻($t+1$)的相对信念仅依赖于历史 h^t 和局中人 i 在时刻 t 的行动:

$$\mu^*(\theta_i|(\theta_i,\theta_i'),(h^t,a^t))=\mu^*(\theta_i|(\theta_i,\theta_i'),(h^t,\tilde{a}^t))\quad 若\ a_i^t=\tilde{a}_i^t\tag{12.30}$$

读者不妨将 $B(1)$—$B(3)$与 $B^*(1)$—$B^*(3)$试作比较,不难发现它们在形式上完全相同。只不过 $B^*(1)$—$B^*(3)$处理的是相对概率而已。事实上,当 $\mu(\cdot|h^t)$对所有历史 h^t 具有全支撑(即在历史 h^t 条件下无零概率类型),那么条件(B)与(B^*)是一致的。特别地,在两周期博弈中的第一周期,所有类型具有正概率,因此条件(B^*)并没有对条件(B)进行精炼(对于在第二周期末所形成的信念,(B^*)的确精炼了(B),但那些信念是无关紧要的,因为博弈到此业告结束)。在每个局中人至多两个类型的情况,任何历史 h^t 之后显然至多只有一个类型具有零概率,因而相对信念问题不再出现(此时绝对信念也是相对信念),从而再一次地,条件(B^*)与条件(B)相一致。

条件 $B^*(1)$意指,在给定历史 h^t 下,如果类型 θ_i 比类型 θ_i' 有无穷多个可能,并且 $\sigma_i(a_i^t|h^t,\theta_i)>0$,那么在 t 周期观察到行动 a_i^t 之后,θ_i 仍然比 θ_i' 有无穷多可能。类似地,如果两个类型在给定 h^t

下是“等可能的”(即，没有一个类型比其他类型有无限多可能)，并且两个类型都以正概率取行动 a_i^t，那么两个类型仍然是等可能的。这两个结论只要从(12.26)式立即可得。这两种讲法排除了我们在图 12.13 中所标示的信念。

基于这些准备工作，我们将完美 Bayes 均衡的概念也给予延拓：

定义 12.4 具独立类型的不完全信息多阶段博弈的完美延拓 Bayes 均衡(perfect extended Bayesian equilibrium)是满足条件(P)与(B^*)的广义状态(σ,μ^*)。

Fudenberg 与 Tirole 于 1991 年得到如下结论：

定理 12.3 (Fudenberg & Tirole 1991)对于具有独立类型的不完全信息的多阶段博弈，条件(B^*)蕴含了条件(c)，任何满足(c)的状态可以延拓为满足条件(B^*)的广义状态。因此，完美延拓 Bayes 均衡与序贯均衡相一致。

我们略去定理 12.2 与定理 12.3 的证明。

§12.4 颤抖手完美均衡

利用颤抖处理零概率行动，由此引出在一般展开型博弈中的序贯均衡，以对完美 Bayes 均衡进行精炼。本节考虑策略型与代理人策略型(agent-strategic form)中的颤抖手完美概念，今后简单地称颤抖手完美均衡(Trembling-Hand Perfect Equilibrium)为“完美均衡”。这是 Selten 的一个贡献，我们将指出，策略型中的“完美”并非以前在展开型中所强调的子博弈完美。Selten 为了排除子博弈不完美均衡而引进了代理人策略型。

对于策略型博弈来说，如何应用“颤抖”手段呢？以一个最简单

的特例来说明，对局中人 i 的某个策略 σ_i，我们可以对任意固定的正数 $\varepsilon>0$，构造一个混合策略 σ_i^ε，使得对一切 $s_i\in S_i$，有 $\sigma_i^\varepsilon(s_i)\geqslant\varepsilon$，如果当 ε 趋于零时，σ_i^ε 收敛于 σ_i，显然 σ_i^ε 系列可以看作局中人 i 取 σ_i 时发生的颤抖。如果上述讲法对一切局中人 $i(i=1,2,\cdots,n)$ 成立。那么策略剖面系列 σ^ε 可以视作所有局中人的策略剖面 σ 的"颤抖"。现在可以这样地认识问题：局中人想采取策略剖面 σ 时可能发生一系列的颤抖，或者解释为局中人在取 σ 时发生了"小错误"。如果在给定 $\varepsilon>0$ 的条件下，也就是说在给定所有局中人都有可能犯一定的小错误的条件下，每个局中人 i 采取的 σ_i^ε 是他的最优策略，且当 $\varepsilon\to0$ 时，策略剖面 $\sigma^\varepsilon\to\sigma$，那么 σ 成为合理预测的均衡结局看来是件顺理成章之事。我们称在给定 ε 条件下的均衡 σ^ε 为"ε-约束均衡"，$\varepsilon\to0$ 后所得到的极限就是我们即将定义的完美均衡。正式的定义如下：

定义 12.5 策略型博弈的"ε-约束均衡"是满足下述条件的完全混合策略剖面 σ^ε：对任意给定的 $\varepsilon>0$、对每一个局中人 $i(i=1,2,\cdots,n)$ 及某些 $\{\varepsilon(s_i)\}(s_i\in S_i,i=1,2,\cdots,n)$，其中 $0<\varepsilon(s_i)<\varepsilon$，在条件 $\sigma_i(s_i)\geqslant\varepsilon(s_i)$ 对所有 s_i 成立的条件下，σ_i^ε 是如下表达式的解：$\max\limits_{\sigma_i}u_i(\sigma_i,\sigma_{-i}^\varepsilon)$（或写作 $\sigma_i^\varepsilon\in\arg\max\limits_{\sigma_i}u_i(\sigma_i,\sigma_{-i}^\varepsilon)$）。

完美均衡是 ε-约束均衡 σ^ε 当 ε 趋于零时的任何极限。

定义 12.5 指出，完美均衡实际上是一些约束博弈序列的 Nash 均衡的极限。由在第三章中曾提出的标准的闭图讨论容易证明任何完美均衡是无约束博弈 Nash 均衡。对于给定的 $\{\varepsilon(s_i)\}$，通常地，约束均衡存在，其存在性定理几乎完全类似于(混合)Nash 均衡存在性定理，在这里略去。于是，对任何 $\{\varepsilon(s_i)\}$ 序列，总存在着相应的约束均衡序列。由于所考虑的策略空间具有紧致性，该系列具有收敛子序列，该收敛子序列所收敛的极限显然(根据定义 12.3)就是完美均衡，这样，事实上我们证明了完美均衡的存在性。

我们引进“ε-约束均衡”作为策略型博弈的 Nash 均衡(由定义 12.3 知,它是 ε-约束均衡的收敛子序列极限)的颤抖,现在我们必须关心此类颤抖是否为对 Nash 均衡集的精炼,或者说,颤动怎样地有助于精炼 Nash 均衡。最好的办法是用例子来解释,考虑图 12.14 所示的博弈。

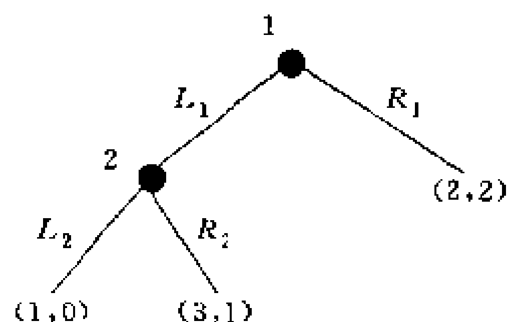


图12.14

这是 Selten 用来启发出子博弈完美均衡的例子。将图 12.14 从展开型改写为策略型(见图 12.15)。

		局中人 2	
		L_2	R_2
局中人 1	L_1	1,0	3,1
	R_1	2,2	2,2

图 12.15

易知,该博弈有两个纯策略 Nash 均衡:(L_1,R_2)与(R_1,L_2)。这里(R_1,L_2)不可能是一系列约束均衡的极限,因为在这个 Nash 均衡中,局中人 1 取策略 R_1,倘若他发生“颤抖”,也就是他“犯了一个小小的错误”而以正概率取行动 L_1,那么局中人 2 赋予尽可能多的权(概率)取 R_2。因此,当 $\varepsilon\to 0$,即局中人 1 取策略 R_1 的概率随之“颤抖”地趋于 0 时,由于颤动的概率为正,局中人必然一直取 R_2,该情况倘若有极限,其结局不可能是(R_1,L_2)。但是注意到 Nash 均衡(L_1,R_2)是一个颤动手完美均衡,因为在这个 Nash 均衡

中，局中人 1 取策略 L_1，如果他发生颤抖而以正概率取行动 R_1 时，考虑到对于局中人 2 来说，策略 R_2 至少是个（弱）优策略，因此约束均衡系列中至少有一个子序列，当 $\epsilon\to 0$ 时（颤抖幅度越来越小，或犯小错误可能性越来越小）趋于（L_1，R_2）。博弈存在着两个纯策略 Nash 均衡，以哪一个预测较合理一些呢？我们的结论是，由于（L_1，R_2）是完美均衡而（R_1，L_2）不是完美均衡，因此也许取（L_1，R_2）作为预测结局更为合理。这个例子显示了颤抖有助于精炼 Nash 均衡集。

现在用一句相当通俗的语言总结定义 12.3，它指出，局中人可能颤抖（或犯错误），当考虑到对手的颤抖时，约束策略应当是局中人的最优策略。Selten 对于完美均衡引入的第二个定义并不像定义 12.3 那样明晰地引入极小颤抖以反映出这句总结性语言，它以对手的扰乱策略（perturbed strategies）σ^m_{-i}来取代考虑对手的“颤抖”，无疑在本质上是一致的。

定义 12.6 策略型博弈的策略剖面 σ 称为完美均衡，如果存在一系列完全混合策略剖面 $\sigma^m\to\sigma(m\to\infty)$，使得对所有 i，$u_i(\sigma_i,\sigma^m_{-i})\geqslant u_i(s_i,\sigma^m_{-i})$ 对一切 $s_i\in S_i$ 成立。

从表面上看，定义 12.6 与定义 12.5 似乎模样很相像，它们之间的区别究竟是什么呢？定义 12.6 中的策略 σ_i 只是关于某些序列 σ^m_{-i}而不必是关于所有收敛于 σ_{-i}的序列的最优反应。而在定义 12.5 中，叙述似乎是对所有 ϵ-约束而言，但结论仅需 σ 是某些序列的约束均衡的极限。因此我们在定义 12.3 中指出的是“任意极限”。如果对一切 $\epsilon>0$，ϵ-约束均衡 σ^ϵ 构成的集合中任意一个聚点都称为完美均衡，由此我们也可意识到完美均衡可以不止一个。两个定义的一致形式是要求 σ_i 是关于收敛于 σ_{-i}的 σ^m_{-i}的最优反应，这个一致形式将产生“真正完美均衡”（truly perfect equilibrium）概念，这是一个更为人们所需要的概念。然而，对某些博弈，真正完

美均衡未必存在。

下面我们将要介绍由 Myerson 于 1978 年提出的完美均衡第三种定义，它并不涉及通常的最优化。

定义 12.7 策略性的策略剖面 σ^{ϵ} 是一个 ϵ-完美均衡，如果

(1)σ^{ϵ} 是一个完全混合策略剖面；

(2)对所有 $i(i=1,2,\cdots,n)$ 与任何 s_i，如果存在 s_i' 满足

$$u_i(s_i,\sigma_{-i}^{\epsilon})<u_i(s_i',\sigma_{-i}^{\epsilon})$$

那么

$$\sigma_i^{\epsilon}(s_i)<\epsilon$$

又，

一个完美均衡 σ 是对某些收敛于 0 的正数序列 ϵ 的 ϵ-完美均衡剖面 σ^{ϵ} 的任意极限。

定义 12.7 中所规定的 ϵ-完美均衡，从不等式方面来看，显然 σ_i^{ϵ}不是关于 σ_{-i}^{ϵ}的最优反应，就是说，我们并没有要求局中人 i 关于其对手的策略采取通常的最优对策，但定义要求对于那些并非最优反应的策略上，混合纯策略 σ_i^{ϵ}只能置小于 ϵ 的权。换句话说，假如纯策略 s_i 不是局中人 i 关于 σ_{-i}^{ϵ}的最优策略，那么 s_i 在混合策略中只占相当小(小于 ϵ)的份量。

迄今为止，我们已经赋予完美均衡三种在提法上有所不同的定义，人们关心的是，这三种定义是否等价，下面的定理回答了这个问题：

定理 12.4 关于完美均衡的三种定义，定义 12.5、定义 12.6 与定义 12.7 是等价的。

证明：沿下述顺序证明本定理：定义 12.5→定义 12.7→定义 12.6→定义 12.5。

定义 12.5→定义 12.7：

考虑 σ^{ϵ} 是定义 12.5 中所定义的 ϵ-约束均衡，即 σ^{ϵ} 满足：对每

一个局中人 i，在条件 $\sigma_i(s_i) \geq \varepsilon(s_i)$ 下（对所有 s_i，且 $0<\varepsilon(s_i)<\varepsilon$）

$$\sigma_i^{\varepsilon} \in \arg\max_{\sigma_i} u_i(\sigma_i, \sigma_{-i}^{\varepsilon})$$

由此，对任何 s_i，如果存在另一个 $s_i' \neq s_i$，使得 $u_i(s_i, \sigma_{-i}^{\varepsilon}) < u_i(s_i', \sigma_{-i}^{\varepsilon})$ 的话，那么只有在 $\sigma_i(s_i) < \varepsilon(s_i) < \varepsilon$ 下才有可能，这实际上就是定义 12.7 中对 ε-完美均衡所要求的。因此凡满足定义 12.5 的 σ^{ε} 必定满足定义 12.7。

定义 12.7→定义 12.6：

假定 σ 满足定义 12.7，那么存在系列 $\sigma^{\varepsilon} \to \sigma$，并且存在一个常数 $d>0$，对于 σ_i 的支撑中的每一个 s_i 成立 $\sigma_i^{\varepsilon}(s_i)>d$（该 σ_i 赋予 s_i 正概率，则称 s_i 为 σ_i 的支撑）。由定义，σ_i 的支撑中的每一个 s_i 必定是 $\sigma_{-i}^{\varepsilon}$ 的最优反应，这恰好符合定义 12.6 的叙述。

定义 12.6→定义 12.5：

假如 σ 满足定义 12.6，且令 σ^m 为收敛于 σ 的一系列定义中所规定的完全混合策略剖面。对于 σ_i 的支撑集中的 s_i，令 $\varepsilon^m(s_i)=1/m$，对于不是 σ_i 的支撑集中的 s_i，则令 $\varepsilon^m(s_i)=\sigma_i^m(s_i)$（注意：尽管 s_i 不属 σ_i 的支撑集，但由于 σ^m 为完全混合策略剖面，必有 $\sigma_i^m(s_i)>0$）。现考虑规划 $\{\max_{\sigma_i} u_i(\sigma_i, \sigma_{-i}^{\varepsilon})$，在 $\sigma_i(s_i) \geq \varepsilon^m(s_i)$ 对所有 $s_i \in S_i$ 成立的条件下$\}$，由定义 12.6 的假设，σ_i 是关于 $\sigma_{-i}^{\varepsilon}$ 的最优反应，相应的 ε-约束均衡之一，σ^{ε}，在 s_i 不属于 σ_i 的支撑集时有 $\sigma_i^{\varepsilon}(s_i)=\varepsilon^m(s_i)$，若 s_i 属于 σ_i 的支撑集则有 $\sigma_i^{\varepsilon}(s_i)=\sigma_i(s_i)$。令 $\varepsilon^m=\max\{\varepsilon^m(s_i)\}$，当 m 趋于无限时，ε^m 趋于 0。显然，此时定义 12.5 得到满足。

Selten 提出，策略型中的完美不是完全令人满意的。现考虑图 12.16。该博弈的策略型表示见图 12.17。

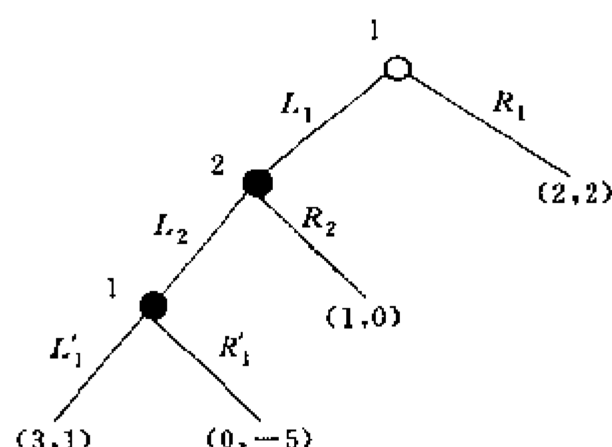


图12.16

		局中人 2	
		L_2	R_2
	R_1	2,2	2,2
局中人 1	L_1,L_1'	3,1	1,0
	L_1,R_1'	0,−5	1,0

图 12.17

由后退归纳法得到唯一的子博弈完美均衡(L_1,L_2,L_1')。从图12.7知,(R_1,R_2)为 Nash 均衡,由此(R_1,R_2,R_1')也是博弈的 Nash 均衡,且它不是子博弈完美。观察一下(R_1,R_2,R_1')是否具颤抖的均衡的极限呢?考虑图 12.7 策略型,显然这个策略型表示式是经过缩减或压缩的,按图 12.16 原意,局中人 1 的策略中 R_1 应分为(R_1,L_1')与(R_1,R_1')两种,不过它们与 L_2 或 R_2 匹配所相应的盈利都是(2,2),故压缩为图 12.17 形式。现在令局中人 1 以概率 ε^2 取(L_1,L_1')且以概率 ε 取(L_1,R_1')。在局中人 1 已经取行动 L_1 的条件下他取 R_1'的条件概率为 $\varepsilon/(\varepsilon+\varepsilon^2)\simeq 1$(当 ε 相当小时),因此为极大化自己的盈利,局中人 2 将置尽可能多的概率权于取 R_2。也就是说,当局中人 1 在取(R_1,R_1')时发生颤抖,颤抖的均衡极限显然仍是(R_1,R_2,R_1')。这个事实足以指出策略型中的完美均

衡(R_1, R_2, R_1')不是令人满意的，因为它连子博弈完美这一点都达不到。问题出在什么地方呢？关键在于策略型颤抖允许在局中人的颤抖与他在后面的信息集上的行动之间存在相关。例如，以上面这个例子来看，假如局中人 1 颤抖而取 L_1，那么他在下一次的行动很可能取 R_1'而不取 L_1'。正是因为局中人的颤抖可能是相关的，因此子博弈完美概念太"强"了一些，因为子博弈完美的前提是子博弈中的合理行动仅依赖于该子博弈，而不管该子博弈是整个博弈树还是代之于仅当某些局中人 i 偏离较长博弈中的(完美)均衡所到达的子博弈。如果我们按字面取颤抖意思，这个前提可能使人非信不可，或者可能使人不信，它依赖于错误如何发生和为什么发生，子博弈完美失去了一部分它的说服力。

1975 年，Selten 修改了他的颤抖手均衡概念以排除相关性，于是也排除了子博弈完美均衡。该修改使用了代理人策略型(agent-strategic form)概念，这个概念将图 12.16 中局中人的两种选择处理成由两个独立颤抖的不同局中人的所为。

更确切地，在代理人策略型中，每个信息集由不同的"代理人"行动，我们可以视作一个委托人将每个信息集上的行动委托给某个代理人实施，不同信息集上委托的不同代理人其行动是独立的。并且在信息集 h 行动的代理人在终点结上有着与原博弈中在 h 行动的局中人 $i(h)$相同的盈利。显然，展开型博弈的代理人策略形式中的颤抖手完美均衡是对应的展开型中的颤抖手完美均衡。因此完美均衡各种定义之间的等价性等问题无疑可以延拓到代理人策略型的完美。今后，对于"完美均衡"，我们将意指"代理人策略型中的颤抖手完美均衡"(与"策略型完美均衡"相对立，它允许相关颤抖贯穿同一局中人的信息集)。我们将图 12.16 所示的展开型博弈描绘为与之相联的代理人策略型，以阐述代理人策略型均衡，见图12.18。

	L_1'	R_1'
L_2	3,1	0,−5
R_2	1,0	1,0

L_1

	L_1'	R_1'
L_2	2,2	2,2
R_2	2,2	2,2

R_1

图 12.18

在图 12.16 所示的展开型博弈中，局中人 1 有行动的信息集有两个，在初始结（信息集 h^0）他委托的“第一个代理人”选择行动 L_1 或者 R_1，这在图 2.18 中体现为第一代理人“负责”选择两个矩阵中的某一个：左边的或者右边的。局中人 1 在第二次行动的信息集上委托“第二个代理人”“全权处理”，这相当该代理人在图 12.18 中负责选择盈利矩阵中的列。根据规定，这两个代理人在终点结上的盈利应与局中人 1 在原博弈中相应终点结上的盈利一样，因此各盈利均标绘在图 12.18 中盈利矩阵的各个结局中。第一代理人与第二代理人互相独立地采取行动，从而“避免”了贯穿局中人 1 的两个信息集上的相关颤抖。由于两个代理人独立选择策略，考虑 $\{R_1, R_2, R_1'\}$，当第一个代理人犯错误或发生颤抖而以某小概率选择 L_1 时，图 12.18 左边的博弈将告诉我们，第二个代理人的 R_1' 弱劣于 L_1'，因此它不可能置较大概率于 R_1'，考虑到这一点，局中人 2 将置较大概率于 L_2 而不是 R_2，不难发现 $\{R_1, R_2, R_1'\}$ 此时不可能成为颤抖均衡的极限，事实上，代理人策略型完美均衡排除了非子博弈完美 Nash 均衡 $\{R_1, R_2, R_1'\}$。

完美均衡（即代理人策略型的颤动手完美均衡）是序贯均衡，为什么呢？只需要注意到先将定义 12.5—定义 12.7 延拓到代理人策略型（不妨将延拓而成的定义仍然称作定义 12.5、定义 12.6 与定义 12.7，只不过原来定义中的“策略型”应改为“代理人策略型”），然而从定义 12.6 就可以清晰地看到这一点。根据定义中所提及的完美均衡的结构，策略 σ 是完全混合策略 σ^n 的极限。为了

得到序贯均衡，人们必须构造信念 μ 使得(σ,μ)为一致的并且在给定信念 μ 下 σ 是序贯理性的。因为 σ^m 是完全混合的，而且具有此策略形式的任何展开型的信息集上与 σ^m 相联的信念是通过 Bayes 法则唯一地确定的。于是仅需取 μ^m 的一个收敛子序列(不妨令该子序列即为 μ^m)的极限作为我们企图构造的信念 μ。根据构造过程，显然(σ,μ)是一致状态。再由定义 12.6 知，对于每一个局中人 i，对所有 $s_i\in S_i$ 成立，$u_i(\sigma_i,\sigma_{-i}^m)\geqslant\mu_i(s_i,\sigma_{-i}^m)$，由一步偏离准则，必有 $u_{i(h)}((\sigma_i,\sigma_{-i}^m)\mid h,\mu^m(h))\geqslant u_{i(h)}((s_{i(h)},\sigma_{-i(h)}^m)\mid h,\mu^m(h))$对一切 $s_{i(h)}$ 成立，因为盈利函数的连续性，令 $m\to\infty$ 即得状态(σ,μ)满足序贯有理条件。至此，我们证明了完美均衡为序贯均衡。

但是，序贯均衡不一定是完美均衡。见图 12.19 所示博弈：

		局中人 2	
		L	R
局中人 1	U	1,1	0,0
	D	0,0	0,0

图 12.19

在这个博弈中，(D,R)是个 Nash 均衡，由于博弈的进行是局中人同时行动，容易验证(D,R)是序贯均衡，但是(D,R)不是完美均衡，设想局中人 1 若发生颤抖，以概率 ε 取 U，那么局中人 2 一定以尽可能大的概率取 L。

但是上述例子是比较特殊的，因为对于 Nash 均衡(D,R)来说，两个局中人对于取均衡策略或取非法均衡策略显然感到无所谓(因为取非均衡策略不会减少他们的盈利)。一旦由于盈利的小小扰动而使这种“无所谓”遭到破坏，序贯均衡集合与完美均衡集合就相同，这是 Kreps 与 Wilson 于 1982 年证明了的结果。

定理 12.5 (Selten 1975)在有限博弈中，至少存在一个完美均衡。

定理 12.6 (Kreps & Wilson 1982)完美均衡是序贯均衡，但其逆不真。然而，对普通的博弈，这两个概念是一致的。

所谓博弈是普通的，"普通性"有严格的定义，这里我们仅指出，此类普通的博弈是对几乎所有的博弈而言。

§12.5 适度均衡

1978 年 Myerson 基于如下思想研究了一个扰动博弈，这个思想是，对那些不太损害自己的行动，局中人更有可能犯错误(或颤抖)，也就是说，偏离均衡行为的概率与为此付出的代价是逆向相联的，代价越高，偏离的概率就越小。这很符合日常生活中人们处理事情的常态，对那些一旦发生错误就会付出沉重代价的行动，局中人会小心翼翼地处理；对那些无需太多代价的行为，局中人常因为无所谓而犯一些小错误。Myerson 的扰动博弈里，考虑了比起最优行动，采取第二最优行动的概率至多是前者的 ϵ 倍，类似地，赋予第三最优行动至多 ϵ 倍的第二最优行动概率，依此类推。

现在我们来阐述由此引出的适度均衡(proper equilibrium)概念，考虑图 12.20 所示博弈，这个例子属于 Myerson。

		局中人 2		
		L	M	R
局中人 1	U	1,1	0,0	−9,−9
	M	0,0	0,0	−7,−7
	D	−9,−9	−7,−7	−7,−7

图 12.20

将图 12.20 与图 12.19 相比较，只不过为两个局中人各添加了弱劣策略 D 与 R。显然该博弈有三个纯策略 Nash 均衡(U,L)、(M,M)和(D,R)。由于 D 与 R 分别是弱劣策略，因此当其他局中人颤动而以小概率偏离 D(或 R)时，局中人肯定以尽可能大的概

率选择另外的策略，因此(D,R)不是完美的。但是与图 12.19 相比，此时(M,M)却是完美的。这又是为什么呢？考虑一个完全混合策略：每个局中人都以$(1-2\varepsilon)$概率取 M，而对另外两个策略各取概率 ε。由于盈利矩阵及混合策略关于两个局中人均对称，仅需考虑局中人 1 的盈利。局中人 1 在给定局中人 2 的混合策略$(\varepsilon,1-2\varepsilon,\varepsilon)$之后各个纯策略的期望盈利分别等于

$$u_1(U,\sigma_2^\varepsilon)=\varepsilon-9\varepsilon=-8\varepsilon$$

$$u_1(M,\sigma_2^\varepsilon)=-7\varepsilon$$

$$u_1(D,\sigma_2^\varepsilon)=-9\varepsilon-7(1-2\varepsilon)-7\varepsilon=-2\varepsilon-7$$

注意到$\sigma_1^\varepsilon(U)=\sigma_1^\varepsilon(D)=\varepsilon$，仅观察局中人 1 取 M 时盈利，只要正数 ε 适当地小，局中人 1 取 U 及取 D 时的盈利总是小于(-7ε)。根据定义 12.7，$\sigma^\varepsilon=((\varepsilon,1-2\varepsilon,\varepsilon),(\varepsilon,1-2\varepsilon,\varepsilon))$是 ε-完美均衡。且当 $\varepsilon\to 0$ 时 $\sigma^\varepsilon\to(M,M)$，故$(M,M)$为完美均衡。以 Myerson 的观点，对于弱劣策略 D 至少不能赋予与 U 一样多的权(不妨认为 D 是第三最好行动，而 U 至少是第二最好行动)。假设局中人 1 赋予 U 以概率 ε，以 Myerson 的观点，局中人 1 应赋予 D 以概率 ε^2，我们来考虑局中人 2 的各纯策略盈利，局中人 2 取 M 时得$(-7\varepsilon^2)$，而取 L 时则得盈利 $\varepsilon-9\varepsilon^2$，当 ε 比较小时，显然有$(\varepsilon-9\varepsilon^2)>(-7\varepsilon^2)$，即局中人 2 取 L 优于取 M。因此，如果在策略型中，除了完美均衡条件外再希望加上 Myerson 的想法，那么(M,M)就不符合条件了，依照下面定义 12.8 的说法，(M,M)不是适度均衡。

定义 12.8 ε-适度均衡是一个完全混合策略剖面 σ^ε，它满足如下条件：如果对一切 i，$u_i(s_i,\sigma_{-i}^\varepsilon)<u_i(s_i',\sigma_{-i}^\varepsilon)$，$\forall\, s_i,s_i'\in S_i$，那么必有 $\sigma_i^\varepsilon(s_i)\leqslant\varepsilon\sigma_i^\varepsilon(s_i')$ 。

如果当 ε 趋于 0 时，ε-适度均衡 σ^ε 的任意收敛子序列的极限 σ 称为适度均衡。

定理 12.7 (Myerson 1978)所有有限策略型博弈都一定有

适度均衡。

定理 12.7 实际上是有限策略型博弈中适度均衡的存在性定理，我们在这里略去了它的证明。在图 12.20 所示的博弈中，一共只有三个 Nash 均衡且都是纯策略的。(D,R)不是完美的，完美的(M,M)不是适度的，因此，(U,L)是该策略型博弈的唯一适度均衡。其实，真正地验证(U,L)是适度均衡也不难，对 $0<\varepsilon<1$，令每个局中人的混合策略为$(1-\varepsilon-\varepsilon^2,\varepsilon,\varepsilon^2)$，读者可以证明 σ^ε 是 ε-适度均衡，而且当 $\varepsilon\to 0$ 时，σ^ε 趋于纯策略(U,L)。

正如前述，适度均衡的引进与日常人们的常规思想相一致，因此很容易被人们所接受，同时，它所具有的良好性质使人们更体会出适度均衡的合理性。

首先，无需借用代理人策略型，适度均衡产生后退归纳效果。这是因为关于颤抖方面的特殊要求保证了局中人在非均衡路径上最优地行动。为阐述这一点，回想图 12.16，只要局中人 2 颤抖，显然(L_1,R_1')劣于(L_1,L_1')。因此，如果局中人 1 的第二个信息集到达的话，局中人 1 必定将几乎所有的概率置于 L_1'。也就是说，(R_1,R_2)Nash 均衡，在非均衡路径上根据适度均衡的思想我们将取(L_1,L_1')，于是得到的是后退归纳解(L_1,L_2,L_1')。

其次，Kahlberg 与 Mertens 于 1986 年证明了“策略型博弈中的每一个适度均衡是与给定策略型相应的每一个展开型博弈中的序贯均衡”。同时，Kohlberg 与 Mertens 也注意到策略型博弈的适度均衡不必是与该策略型相联的每一个展开型的(代理人策略型中的)颤抖手完美均衡。为阐述这个问题，考虑最简单的单一决策者问题，含有三个纯策略，见图 12.21。

在图 12.21 所示的策略型战争中，(L,r)与后退归纳解一致，它是适度的，但是它在博弈树中不是(代理人策略型)完美的；假如局中人的第二个代理人颤抖，它的第一个代理人宁可取 R 为上。

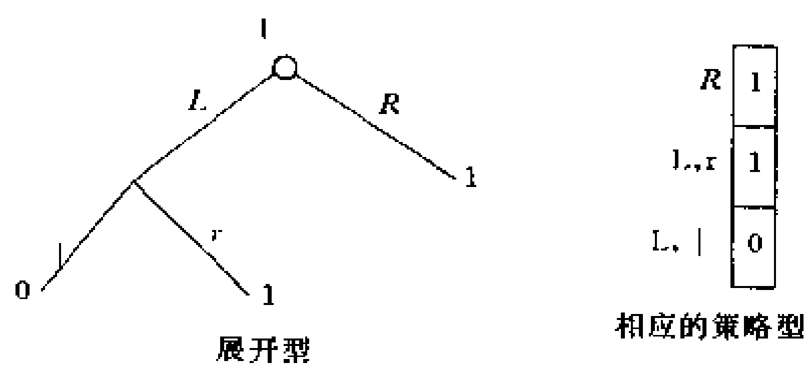


图12.21

第十三章 信号博弈

§13.1 信号博弈中的完美 Bayes 均衡

不完全信息动态博弈最简单的例子之一是信号博弈，它包含两个局中人：发送者（记为 S）与接收者（记为 R），因为是动态的，博弈时序规定如下：

（1）自然按照概率分布 $p(t_i)$ 为发送者 S 从一个可行类型空间 $T=\{t_1,\cdots,t_I\}$ 中选取类型 t_i，其中 $p(t_i)>0$ 对每一 i 成立，且 $p(t_1)+\cdots+p(t_I)=1$。

（2）发送者 S 观察到 t_i 后，从一个可行信号集 $M=\{m_1,\cdots,m_J\}$ 中选取一个信号 m_j。

（3）接收者 R 观察到信号 m_j（注意，不是观察到 t_i），然后从可行行动集 $A=\{a_1,\cdots,a_K\}$ 中选取行动 a_k。

（4）S 与 R 的盈利（或效用）函数分别为 $U_s(t_i,m_j,a_k)$ 与 $U_R(t_i,m_j,a_k)$。

这里，我们简单地将类型空间、可行信号集与可行行动集定义为有限集合，在实际应用中，它们常常表现为连续统的区间，显然，此时可行信号集依赖于类型空间，而可行行动集则依赖于发送者发出的信号。

现在考虑图 13.1 所展示的博弈。

这是一个简单的抽象信号博弈，其中 N 表示自然，$T=$

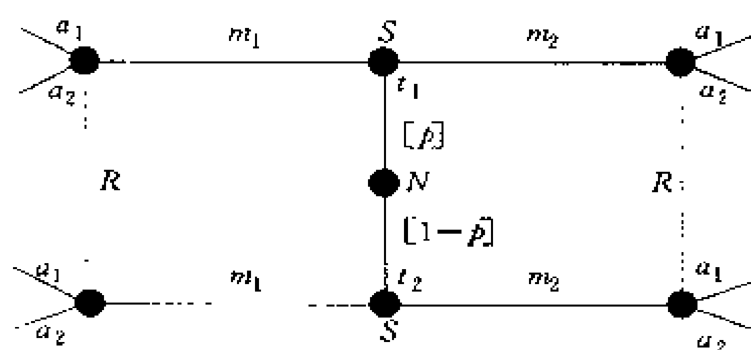


图13.1

$\{t_1,t_2\}$，$M=\{m_1,m_2\}$，$A=\{a_1,a_2\}$，图中$[p]$及$[1-p]$表示自然选择类型时的概率分布。注意，这个博弈依时序应先从 N 开始行动，但我们不能将博弈树的开头表示为图 13.2 所示。

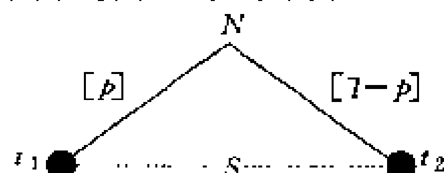


图13.2

图 13.2 的表示是错误的，因为在图 13.2 中，发送者不知道自己属何种类型，但在事实上，发送者知道自己的类型。而正确的图 13.1 告诉我们，只有接收者 R 不知道 S 的类型，他只能依据发送者发出的信号来选择自己的行动。

在任何博弈中，局中人的策略其实是一个完整的行动计划，在局中人可能被要求采取行动的每一个偶然场合，一个策略确定了该场合下的一个可行行动。在信号博弈中，发送者的纯策略是根据自然抽取的可能类型来选取相应的信号，因此，信号可视作类型 t 的函数 $m(t_i)$。而接收者的纯策略是信号的函数 $a(m_j)$，即根据观察到的发送者发出的信号确定自己的行动。在图 13.1 的信号博弈中，S 与 R 各有四个纯策略。

发送者 S 的纯策略：

$S(1)$　若自然抽取 t_1 则取 m_1，若自然抽取 t_2 仍取 m_1；

$S(2)$　若自然抽取 t_1 则取 m_1，若自然抽取 t_2 则取 m_2；

$S(3)$　若自然抽取 t_1 则取 m_2，若自然抽取 t_2 则取 m_1；

$S(4)$　若自然抽取 t_1 则取 m_2，若自然抽取 t_2 仍取 m_2。

接收者 R 的纯策略：

$R(1)$　若 S 发出 m_1 则取 a_1，若 S 发出 m_2 仍取 a_1；

$R(2)$　若 S 发出 m_1 则取 a_1，若 S 发出 m_2 则取 a_2；

$R(3)$　若 S 发出 m_1 则取 a_2，若 S 发出 m_2 则取 a_1；

$R(4)$　若 S 发出 m_1 则取 a_2，若 S 发出 m_2 仍取 a_2。

注意到一个有趣的事实，发送者 S 的纯策略中的 $S(1)$ 与 $S(4)$ 有一个特点，对于"自然"抽取的不同类型，S 选取相同的信号，具有这类特点的策略称为共用(pooling)策略。至于 $S(2)$ 与 $S(3)$，由于对不同的类型发出不同的信号，故称为分离(separating)策略。由于在这个简单情况中各种集合只有两个元素，由此局中人的纯策略也只有共用与分离这两种，假如类型空间的元素多于两个，那么就有部分共用(partially pooling)或半分离(semi-separating)策略。实际上各种类型分为不同的组，对于给定的类型组中所有类型，发送者发出相同的信号，而对于不同组的类型则发生不同的信号。当然，图 13.1 博弈只分共用与分离这两种策略是针对纯策略来说的，在这两种类型的展开型博弈中，也存在着混合策略，例如，若自然抽取 t_1 时 S 取 m_1，而当自然抽取类型 t_2 时，S 则在 m_1 与 m_2 这两个信号中随机选择，这样的混合策略，我们称为混杂(hybrid)策略。

显然，图 13.1 中的信号博弈是不完全信息动态博弈，我们将正式定义此类信号博弈的完美 Bayes 均衡。事实上我们需要将第十二章中的 R_1—R_3(乃至 R_5)在信号博弈中具体化、精确化。为使事情简单化，不妨先仍限于纯策略讨论。

必须清楚如下事实，尽管先由自然为发送者选择类型，但是发送者清楚自己的类型，这正像自然为每个人选择性别，于是人类的生育，生男生女是有随机规律的，但是每一个人完全明白自己的性

别。也就是说，发送者在选择发送什么样的信号时，他知道博弈的历史，于是他自己的选择一定发生于单结信息集（对于自然可能抽取的每一种类型，一定存在这样的单结信息集）。既然信息集仅为单结，因此 R_1 对发送者是平凡的。由图 13.1 直接看出，接收者 R 在观察到发送出来的信号，比方说观察到 m_1，他仍不知道发送者的类型，因此接收者 R 的选择行动发生于图 13.1 左边的非单一信息集。于是 R_1 对接收者是有意义的。在这里我们将它记为 $SR(1)$ 以表示信号博弈中的 R_1：

$SR(1)$　在观察到来自 M 的任何信息 m_j 后，接收者 R 关于发出信号 m_j 的发送者属于哪一种类型必须有一个信念。这个信念应是一个概率分布，由于信念是基于信号 m_j 的，因此可以记作 $\mu(t_i|m_j)$，其中对 T 中的每一个 t_i，应有 $\mu(t_i|m_j)\geqslant 0$，且

$$\sum_{t_i\in T}\mu(t_i|m_j)=1 \tag{13.1}$$

$\mu(t_i|m_j)$ 无非是接收者根据接收到的信号判断发送者属于哪一种类型的概率或可能性，因为接收者一旦完全知道 S 的类型后，他将可以选择使自己盈利极大化的行动。可惜，他只能观察到 m_j 而无法观察到 t_i，因此只能依据信念 $\mu(t_i|m_j)$ 来计算自己的期望盈利。这就是涉及到 R_2 对接收者提出的要求，简记为 $SR(2R)$：

$SR(2R)$　对于 M 中的每一个 m_j，接收者 R 的行动 $a^*(m_j)$ 必须在给定信念 $\mu(t_i|m_j)$ 下极大化自己的期望盈利（或效用）。即，$a^*(m_j)$ 是下述公式的解：

$$\max_{a_k\in A}\sum_{t_i\in T}\mu(t_i|m_j)U_R(t_i,m_j,a_k) \tag{13.2}$$

注意：R_2 并非仅对接收者（在非单一信息集上的行动者）要求极大化期望盈利，而且对发送者（该信息集的前一个行动者，或博弈树较高处的行动者）也要求极大化 S 的盈利，只有这样，才能保证策略至少是均衡。事实上，倘若发送信号 m_j^* 比发送信号 m_j 时

具有更大盈利的话，那么 S 必然会发送 m_j^*。在信号博弈中，发送者有完全信息（已经指出，他具有平凡信念），并且他仅在博弈的开头采取行动，因此 R_2 只是简单地使发送者在给定接收者的策略下采取最优策略，我们用 $SR(2S)$ 来表示之。

$SR(2S)$　对于 T 中的每一个 t_i，发送者发出信号 $m^*(t_i)$ 必须极大化发送者在给定接收者的策略 $a^*(m_j)$ 下的盈利，即 $m^*(t_i)$ 是（13.3）式的解：

$$\max_{m_j \in M} U_s(t_i, m_j, a^*(m_j)) \tag{13.3}$$

最后，如果给定了发送者的策略 $m^*(t_i)$（例如根据 $SR(2S)$ 计算而得），令 T_j 表示发送信号 m_j 的类型集，即，如果 $m^*(t_i)=m_j$，那么 $t_i \in T_j$。倘若 T_j 为非空集合，那么对应于信号 m_j 的信息集显然在均衡路径上；要不，倘若任何类型均不发送信号 m_j，那么对应的信息集当然不在均衡路径上。现在考虑在均衡路径上信号，应用 R_3 于接收者的信念，我们得到：

$SR(3)$　对于 M 中的每一个 m_j，如果存在 T 中的 t_i 使得 $m^*(t_i)=m_j$，那么在对应于 m_j 的信息集上，接收者的信念必定由 Bayes 法则与发送者的策略来确定：

$$\mu(t_i \mid m_j)=\frac{p(t_i)}{\sum_{t_i \in T_j} p(t_i)} \tag{13.4}$$

注：我们这里仅考虑纯策略，因此信念由（13.4）式决定，至于出现混合策略则在稍后实例中进一步阐述。

现在考虑 R_4，由对图 12.5 讨论的情况可知，对于信号博弈而言，R_4 是毫无意义的，或者说是空洞的。因为接收者总是根据观察到的信号建立有关发送者类型的信念。

我们已经可以根据 R_1—R_3 来定义信号博弈中的完美 Bayes 均衡，不过刚才我们将这些条件具体化了。

定义 13.1　信号博弈中的纯策略完美 Bayes 均衡由满足 SR

(1)、$SR(2R)$、$SR(2S)$与 $SR(3)$ 的策略$(m^*(t_i),a^*(m_j))$与信念$\mu(t_i|m_j)$所组成。

如果发送者的策略是共同的，则称完美 Bayes 均衡为共用的；如果发送者的策略是分离的，则相应的完美 Bayes 均衡称为是分离的。

现在我们将图 13.1 中的博弈树赋予盈利向量(见图 13.3)，然后来计算纯策略完美 Bayes 均衡。

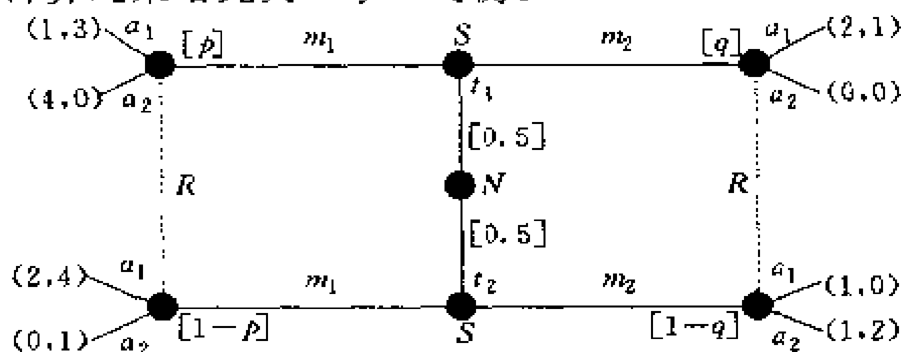


图13.3

在这个具有两个类型、两种信号的博弈中，由于发送者有四个纯策略，因此我们可以从这四个纯策略出发，分别分析该信号博弈的完美 Bayes 均衡。

(1)共用 m_1

发送者的这个纯策略表明，不管发送者的类型是 t_1 还是 t_2，总是发出同样信号 m_1，可以用(m_1,m_1)表示，向量中第一第二个元素分别对应于 t_1、t_2 类型发出的信号。假如博弈存在一个均衡，其中发送者的纯策略是(m_1,m_1)，那么图 13.3 左边接收者的信息集当然在均衡路径上，我们可以通过 $SR(3)$计算得接收者的信念为$(p,1-p)=(0.5,0.5)$。在这个信念下(事实上，在任何其他信念下也一样)接收者在观察到信号 m_1 之后的最佳反应是 a_1，因为从盈利来看，此时 a_2 显然是接收者的严劣策略。从而类型 t_1 与 t_2 的发送者分别获盈利 1 与 2。那么，两种类型的发送者是否都愿选取信号 m_1 呢？这需要确定如果发送者发出信号 m_2，接收者将如何作

出反应以及在相应反应下发送者的盈利究竟如何。试想接收者对信号 m_2 的反应是取行动 a_2，此时类型 t_1 与 t_2 的发送者将分别获盈利 0 与 1，这少于他们发出信号肯定能获得的 1 与 2，就是说，如果接收者对 m_2 的反应是 a_2 的话，不管哪种类型的发送者都不愿发送信号 m_2。如果接收者对于信号 m_2 的反应是采取行动 a_1，那么类型 t_2 的发送者获盈利为 1，少于他发出信号 m_1 后所得盈利 2，但是，类型 t_1 的发送者将由此而获盈利 2，它超过了若发出信号 m_1 后的盈利 1。由此，如果接收者对信号 m_2 的反应是 a_1 的话，类型 t_1 的发送者愿意发出信号 m_2。注意到我们已经假设存在博弈的均衡，其中发送者的策略是 (m_1,m_1)，那么在这样的均衡中，接收者对 m_2 的反应必定是 a_2。于是在这个均衡中，接收者的策略是 (a_1,a_2)，其中第一个元素是对信号 m_1 的反应，第二个元素是对信号 m_2 的反应。要使接收者的策略 (a_1,a_2) 是均衡策略，我们必须计算一下接收者的盈利，显然，我们只需要检查一下当信号为 m_2 时接收者采取 a_1 与 a_2 的期望盈利，由于该信息集上信念为 $(q,1-q)$，由此，接收者取 a_1 的期望盈利为 q，取 a_2 的期望盈利为 $2(1-q)$ 要使接收者在观察到信号 m_2 后不偏离均衡策略 a_2，必须 $2(1-q)\geqslant q$，即 $q\leqslant\frac{2}{3}$。

结论：$[(m_1,m_1),(a_1,a_2),p=0.5,q\leqslant\frac{2}{3}]$ 是博弈的共用完美 Bayes 均衡。

(2)共用 m_2

发送者的纯策略是 (m_2,m_2) 的话，由先验概率也可得 $q=0.5<\frac{2}{3}$，根据(1)中分析，接收者对于 m_2 的最优反应为 a_2。这样，类型 t_1 的发送者获盈利 0，类型 t_2 的发送者获盈利 1，可是，由于接收者对 m_1 的最优反应是 a_1（不管 p 何值），由此类型 t_1 的发送者若发出信号 m_1 的话可以获盈利 1。可见，发送者的纯策略 (m_2,m_2)

不可能是均衡策略。因为他完全可以偏离该策略而从中提高自己的获益。

(3)分离:类型 t_1 发出信号 m_1,类型 t_2 发出信号 m_2

如果存在一个均衡,其中发送者的纯策略为(m_1,m_2)。那么接收者的两个信息集(图 13.3 左右两边)都在均衡路径上,于是两个信念都可由 Bayes 法则与发送者的策略确定,例如对于 p 而言,

$$p=\mu(t_1|m_1)=p(t_1)/p(t_1)=1$$

同样可得

$$1-q=\mu(t_2|m_2)=p(t_2)/p(t_2)=1$$

即 $q=0$。给定信念 $p=1$,接收者的最优反应仍是 a_1,给定信念 $q=0$,接收者的最优反应是 a_2。因此,两种类型的发送者均获盈利 1。现在我们需要检验一下在给定接收者的策略(a_1,a_2)下,发送者的(m_1,m_2)是否最优呢? 图 13.3 明确地告诉我们,类型 t_2 的发送者如果偏离这个策略不发信号 m_2 而发出 m_1,那么由于接收者反应为 a_1 而使类型 t_2 发送者获盈利 2,这比他发送信号 m_2 时会取得更好的盈利,于是在给定接收者策略(a_1,a_2)下,发送者会有主动偏离(m_1,m_2)的可能,故不存在这样的均衡,其中发送者的策略为(m_1,m_2)。

(4)分离:类型 t_1 发出信号 m_2,类型 t_2 发出信号 m_1

发送者的纯策略为(m_2,m_1),如(3)那样,我们可以利用 Bayes 法则与发送者的策略确定接收者的两个信念:$p=0$ 与 $q=1$。在给定这两个信念的情况下,接收者的最优反应是(a_1,a_1),从而类型 t_1 与 t_2 的发送者均获盈利 2。仍然要看看发送者是否会单方而偏离。如果类型 t_1 发送者偏离而发出信号 m_1,那么由于接收者对 m_1 的反应是 a_1,则使类型 t_1 发送者仅获盈利 1;如果类型 t_2 发送者偏离而发出信号 m_2,接收者反应仍为 a_1,发送者的盈利也为 1。显然,无论发送者属于何种类型,都不可能激励他偏离策略(m_2,m_1)。

结论:[(m_2,m_1),(a_1,a_1),$p=0$,$q=1$]是博弈的分离完美

Bayes 均衡。

§13.2 劳务市场信号博弈

中央电视台《经济半小时》某次节目播放无锡小天鹅集团关于人才引进的话题。某名牌大学博士生向董事长提了一个问题：小天鹅集团引进大量博士、硕士，并给了他们很高的待遇，可能存在的问题在于不是所有博士、硕士都是名副其实的人才，小天鹅集团以学历作为引进人才主要标准之一是否妥当(大意如此)？董事长回答得很干脆，在没有经历过较长时间试用的情况下，只能以学历作为重要选材标准之一。

从信号博弈的角度来看，董事长的谈话精神无疑是对的，本节介绍的劳务市场信号博弈(Job-Market Signaling)将回答这个问题。

本节将 Spence(1973)模型描述为展开型博弈，并研究它的某些完美 Bayes 均衡。博弈的时序如下：

(1)自然确定工人的生产能力 η，它可能为高(H)，也可能为低(L)。$P\{\eta=H\}=q$。

(2)工人认识到自己的能力是高还是低。然后选取一个教育水平(level of education)，$e\geqslant 0$。

(3)两个公司观察到工人的教育(而不是工人的能力)，然后同时对工人开出工资价。

(4)工人接受两个开价中较高的工资，在两个开价一样的情况下用抛钱币的方式来决定。

令 w 为工人所接受的工资。

现在来分别计算工人与公司的盈利函数。对工人来说，为得到教育 e 必须付出一定的成本或代价，这个代价通常与该工人的能力有关，因此记 $c(\eta,e)$ 为成本函数。于是工人的盈利函数应等于

$$U_w=w-c(\eta,e) \tag{13.5}$$

两个公司中没有雇用该工人的那家盈利为 0,而雇用工人的那家公司其盈利为

$$U_F=y(\eta,e)-W \tag{13.6}$$

其中,$y(\eta,e)$表示一个具有教育水平 e 与能力 η 的工人为公司所生产的产值。

对这个模型,我们关心一个完美 Bayes 均衡,在这个均衡中,公司视教育水平 e 为工人努力 η 的一个信号,从而对受较多教育的工人提供较高工资。Spence 于 1973 年的论文中就讲述了这一点。这篇论文指出,即使教育对工人的产值没有什么影响,或者说,即使工人的出产完全独立于 e,为 $y(\eta)$,工资还是随着教育而增加。这正是小天鹅集团董事长的选材之道。1974 年 Spence 推广了这类讨论,允许产值为能力与教育两者的递增函数,显然这个推广了的模型更趋合理一些。于是,工资随教育而增加可以解释为教育对生产能力的影响。我们在本节介绍这个更一般情况的模型。

形式上,总是假定(在教育 e 相同情况下)高能力工人会有较多产量,即 $y(H,e)>y(L,e)$对一切 e 成立。并假定教育 e 的增多并不会减少产出,即 $y_e(\eta,e)\geqslant 0$ 对一切 η 与一切 e 均成立,其中 $y_e(\eta,e)$表示教育水平为 e 且具能力 η 的工人的教育边际产出,实际上是 y 关于 e 的偏导数。

为了可以着手模型的研究,我们需要对教育水平 e 进行量化,众所周知的事实是,从全社会来看,学历越高的工人平均工资也越高,由此,将 e 解释为入学年数似乎是顺理成章之事,但是这种解释有可能引起争论,因为在同一个学校同一个年级的两个人,可以是成绩相差很悬殊(他们的学习成本几乎差不多),或者,名牌大学的博士与普通学校的博士(平均来说)也反映了一定程度能力上的差异(但他们的学习成本也可以视作差不多),由于我们在模型中认为 e 对 η(由此对 y)产生影响,因此,以入学年数作为 e 的度量

的确并不完全精确，从而引起争议。为了回避此类争议，我们不妨将 e 解释为，它测度了所读课程的类型与数目，它也测度了年级（和学校）的质量，同时也测度了在一固定长度学年内所获得的荣誉等等。在这种对 e 的解释下，低能力学生发现在一个给定的学校里达到高年级要困难一些，或者他会发现在一个名牌学校（或者说是更具有竞争性的学校），他想达到相同年级水平也会更困难一些。理想地对 e 作出一些解释之后，前面所提到的完美 Bayes 均衡很能解释一件众所周知的事实：教育水平在博弈中被公司视作工人发出的信号，信号博弈的完美 Bayes 均衡带来的结果是，公司对在给定学校成绩最好的毕业生和从名牌学校毕业的人支付较多的工资，因为他们发送的信号 e 较高。

Spence 模型中有一个关键假设：对于低能力工人来说，教育的边际成本比起高能力工人要来得高一些，即，对于同一个 e，成立

$$C_e(L,e) > C_e(H,e) \tag{13.7}$$

这个假设很容易被人们接受，因为成本函数 C 作为 e 的递增函数，若想从同一个水平 e 提高相同高度 de，那么低能力者显然要比高能力花费更大代价。

我们又遇到了另一个建模方面的困难，在招聘工人上存在着两个公司之间的竞争，这使模型比上一节所述的信号博弈复杂得多。最好的办法是尽可能地将两个公司"合并"为一个。同时我们也注意到一个趋向，假如两个公司激烈地竞争，其后果很可能使他们的盈利随竞争而越来越少。Spence 在建模时注意到了这个事实，因此他作出了又一个假设：公司间的竞争将使期望盈利趋于0。基于这样的假设，在建模时将"合并"的两个公司称之为"市场"，由"市场"作出单一的工资开价 w 并令其相应的盈利函数不是(13.6)式，而是

$$-[y(\eta,e)-w]^2 \tag{13.8}$$

(13.8)式与 Spence 模型的假设有相合之处。由 $SR(2R)$，为极大化盈利，市场将开出一个工资价 w 等于具有教育水平 e 的工人期望产出。市场在观察到工人的 e 之后，给出关于工人能力的信念，于是

$$w(e)=\mu(H|e)y(H,e)+[1-\mu(H|e)]\cdot y(L,e) \qquad (13.9)$$

其中 $\mu(H|e)$ 是市场关于工人能力为 H 的概率评估。在博弈的阶段(3)，两个公司互相针对地标价其目的在于不用求助于虚拟局中人“市场”而达到相同的结果。可是，为了保证公司将总是提供相当于工人的期望产出的工资(即(13.8)式中的 $w(e)$)，我们必须对模型再强加一个假设：在观察到教育水平 e 之后，两个公司对于工人的能力拥有相同的信念 $\mu(H|e)$。这个假设表面上看似乎有点“强加”，但从社会实际效果来看是有一定意义的。长时期的社会实践使公司对工人的能力评估在大部分情况下比较接近。因为 $SR(3)$ 决定两个公司在观察到均衡路径上的 e 之后必须拥有的信念，因此我们的假设实际上是说公司在观察到非均衡路径上的 e 之后也分享共同信念。在这种假设下，得出结论为在任何完美 Bayes 均衡中，两个公司均提出工资如(13.9)式。于是(13.9)式取代了本节所原先提出的存在两个接收者的要求 $SR(2R)$。

在分析完美 Bayes 均衡之前，首先考虑简单化一些的模型，即信息是完全的劳务市场博弈。所谓信息完全，无非是工人的能力 η 成为所有局中人的共同知识。在这种情况下，“市场”(或两个公司竞争的结果)使得具有教育 e 并且具有能力 η 的工人挣得工资 $w(e)=y(\eta,e)$，于是具有能力 η 的工人应当选择适当的 e，以成为下述公式的解：

$$\max_{e}[y(\eta,e)-c(\eta,e)] \qquad (13.10)$$

不妨记(13.10)式的解为 $e^*(\eta)$，令 $w^*(\eta)=y(\eta,e^*(\eta))$，见图 13.4。这个解在下面的分析中常常要用到。

现在我们回到工人的能力 η 假定为私人信息的情况。一旦 η

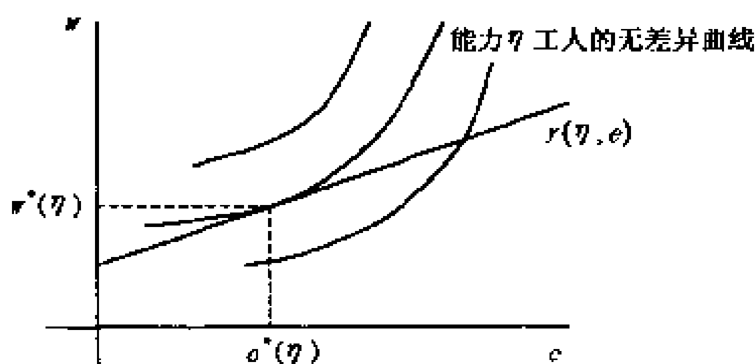


图 13.4

是工人的私人信息，不为公司所知道，那么就有可能低能力工人试图冒充高能力工人接受招聘。这里有两种可能情况，我们用图像加以阐述，图 13.5 就是一种情况。

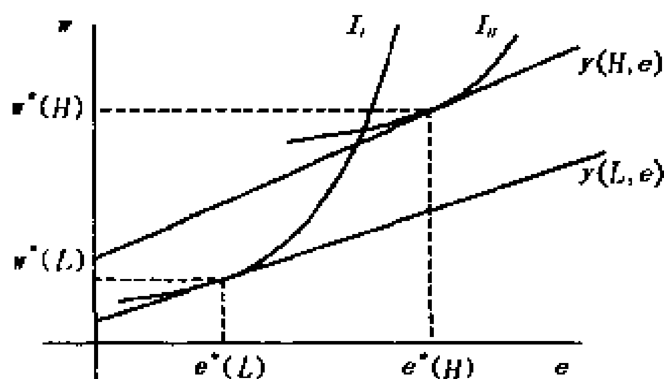


图 13.5

其中 I_L 与 I_H 分别表示低能力与高能力工人的无差异曲线，由(13.7)式，I_L 要比 I_H 更陡一些。图 13.5 描绘了这样的可能，低能力工人由于太昂贵的缘故而不去获得 $e^*(H)$，尽管如果他们这样去做(即去学习并获得 $e^*(H)$)可以哄骗公司会相信他有高能力，从而引起公司付出高工资 $w^*(H)$。由此，在图 13.5 中

$$w^*(L)-C[L,e^*(L)]>w^*(H)-C[L,e^*(H)] \quad (13.11)$$

(13.11)式的实际意思是，一个低能力的工人为寻求 $e^*(H)$ 以获得 $w^*(H)$，但他将花费昂贵的成本 $C[L,e^*(H)]$，从而使他此时的盈利反而小于他以 $e^*(L)$ 获得工资 $w^*(L)$ 时的盈利。或者说，

由于(13.11)式，尽管低能力工人可以哄骗公司，但为了极大化自己的盈利，他不愿意这样做。自然会想到，出现的第二种可能情况是

$$w^*(L)-C[L,e^*(L)]<w^*(H)-C[L,e^*(H)] \quad (13.12)$$

事实上，这是劳务市场信号博弈中更现实与更令人感兴趣的情况，因为在这里低能力的工人羡慕甚至嫉妒高能力工人的教育水平与工资，在不太昂贵的情况下，(13.12)式激励他获得$e^*(H)$，从而使公司相信他并给予工资$w^*(H)$，而他的能力仅仅只是L。当能力η是个连续变量时则应用该情况。

我们仍从共同、分离(与混杂)完美 Bayes 均衡角度进行研究，并仅限于少数几个例子。

1. 共用均衡

在一个共同均衡中，工人类型只有H与L两种，根据“共用”定义，两种类型的工人选择了单一的教育水平，设为e_p。$SR(3)$蕴含着，在观察到e_p之后公司的信念必定是先验信念，$\mu(H|e_p)=q$，且蕴含在观察到e_p之后开出的工资价为

$$w_p=qy(H,e_p)+(1-q)y(L,e_p) \quad (13.13)$$

为完成完美 Bayes 均衡的描述，还需要确定$SR(1)$与$SR(2S)$(注意：工资定价公式意味着已考虑了$SR(2R)$)，对于$SR(1)$，需要确定对于非均衡的教育选择$e\neq e_p$，公司的信念$\mu(H|e)$，然后利用(13.9)计算公司的工资策略，因为在给定信念$\mu(H|e)$下，(13.9)确定的工资使公司的期望盈利极大化。$SR(2S)$则需要验证，不管哪种类型的工人，对于公司的开价$w(e)$作出的最佳反应为$e=e_p$。

存在着一种可能性，公司认为任何非e_p的教育水平e意指工人仅有低能力：$\mu(H|e)=0$对一切$e\neq e_p$成立。这种信念看来似乎令人感到陌生。然而完美 Bayes 均衡的定义中没有什么东西可以将它排除在外，这是因为完美 Bayes 均衡中的R_1—R_3对非均衡路径上的信念没有任何要求，而对于R_4、由于信号博弈的特殊性

——接受着仅对发送的信号作出反应，因此 R_4 在信号博弈中是没有意义的，空的。如果希望以一个例子来阐述这种可能性的话，那么公司认为只有大专秘书专业毕业生才能适应自己招聘的秘书职位可以算是一例。

如果公司的信念为

$$\mu(H|e)=\begin{cases}0 & e\neq e_p\\ q & e=e_p\end{cases} \tag{13.14}$$

那么(13.9)式告诉我们，公司的策略为

$$w(e)=\begin{cases}y(L,e) & e\neq e_p\\ w_p & e=e_p\end{cases} \tag{13.15}$$

在给定公司的策略下，工人选择 e 应为如下公式的解：

$$\max_e[w(e)-c(\eta,e)] \tag{13.16}$$

求(13.16)式的解非常容易，仅需注意 $w(e)$ 的表达式(13.15)式，能力为 η 的工人要么选择 e_p，使自己获得工资 w_p，要么选择 e，使 $[y(L,e)-C(\eta,e)]$ 达到极大化。假如工人的类型为 L(即低能力)，那么这个 e 实际上就是 $e^*(L)$。

请注意，我们描述的共同完美 Bayes 均衡是指，无论高能力还是低能力的工人都选择同一个教育水平 $e=e_p$(即，发送出相同的信号)，因此我们关心的是在任何情况下(我们只想找出一些情况，而不例举所有的情况)，在给定信念(13.14)下这个策略剖面满足 $SR(1)$、$SR(2S)$、$SR(2R)$ 与 $SR(3)$。现在，根据上述分析，能力为 η 的工人关于教育水平 e 的最优选择有两种可能，我们需要知道，在什么样的条件下，两种类型的工人都认为 $e=e_p$ 而不是另一种 e 的选择呢？

先对低能力 $\eta=L$ 类型工人分析，如果 $e_p>e^*(L)$，且假定产值函数 $y(L,e)$ 是 e 的严增函数，那么由于

$$w_p=qy(H,e_p)+(1-q)y(L,e_p)$$

$$\geqslant y(L,e_p)>y(L,e^*(L)) \quad (13.17)$$

从工资收入的角度来看，此时低能力工人愿意选择 e_p，但是由于涉及到教育成本 $C(L,e)$，因此，在上述情况下，仅当

$$w_p - C(L,e_p) \geqslant y(L,e^*(L)) - C(L,e^*(L)) \quad (13.18)$$

时，低能力工人不会选择 $e^*(L)$ 而选择 e_p。对于高能力 $\eta=H$ 的工人在满足条件

$$w_p - C(H,e_p) \geqslant y(L,e^*(L)) - C(H,e^*(L)) \quad (13.19)$$

时，他也不会选择 $e^*(L)$ 而选择 e_p。

结论：当 $e_p>e^*(L)$，(13.18)式与(13.19)式同时成立的条件下，工人的策略为 $[e(L)=e_p, e(H)=e_p]$，公司的信念为(13.14)式中的 $\mu(H|e)$，公司的策略是(13.15)式中的 $w(e)$，它们构成了完美 Bayes 均衡。

继续重申一点，以免读者误解，我们只是讨论了共用完美 Bayes 均衡的一小部分。我们利用验证 $SR(1)$、$SR(2S)$、$SR(2R)$ 与 $SR(3)$，至少得到了某种情况的共用完美 Bayes 均衡。

2. 分离均衡

现在需要转到分离完美 Bayes 均衡。倘若没有嫉妒情况，就是说能力低的工人不羡慕乃至嫉妒高能力工人(在完全信息下的)教育水平与工资，这样，低能力工人不去"冒充"高能力以哄骗公司出高工资。显然，工人作为发送信号者，其自然的分离策略为 $[e(L)=e^*(L), e(H)=e^*(H)]$("实事求是"地发送信号)。由此，$SR(3)$ 确定了在观察到这两个教育水平中任何一个以后公司的信念，显然，$\mu(H|e^*(L))=0$ 和 $\mu(H|e^*(H))=1$。因此由(13.9)式(即 $SR(2R)$)，公司的相应最优策略为 $w(e^*(L))=w^*(L)=y(L,e^*(L))$ 与 $w(e^*(H))=w^*(H)=y(H,e^*(H))$。与讨论共同均衡时一样，事情并非到此为止，我们必须确定如下事项才能完成对分离完美 Bayes 均衡的描述：

(1) 确定非均衡教育选择的公司信念 $\mu(H|e)$，即不同于

$e^*(L)$或$e^*(H)$的e值时的$\mu(H|e)$，一旦确定了$\mu(H|e)$，再由(13.9)式确定公司的策略$w(e)$。

(2)试图证明能力为η的工人对于公司策略$w(e)$的最优反应是选择$e=e^*(\eta)(\eta=H$或$L)$。

实现上述条件的一个信念为

$$\mu(H|e)=\begin{cases}0 & e<e^*(H)\\ 1 & e\geq e^*(H)\end{cases} \tag{13.20}$$

(13.20)式表明，如果工人选择至少为$e^*(H)$的e，公司认为该工人具有高能力；反之，公司则认为他仅具有低能力。例如，公司要求招聘高中以上学历的工人就是(13.20)式的一例。相应地，利用(13.9)式，公司的工资策略是

$$w(e)=\begin{cases}y(L,e) & e<e^*(H)\\ y(H,e) & e\geq e^*(H)\end{cases} \tag{13.21}$$

根据前面的讨论，$e^*(H)$其实是高能力工人关于工资函数$w=y(H,e)$的最优反应，因此在这里显然也是高能力工人关于$w(e)$的最优反应。再考虑低能力工人，已知$w^*(L)$其实是他关于工资函数$w=y(L,e)$的最优反应，而(13.21)式表明，当$e<e^*(H)$时，$W(e)=y(L,e)$，由此无疑当$e<e^*(H)$时，与$e^*(L)$相应的盈利$w^*(L)-C[L,e^*(L)]$是低能力工人可达到的最大盈利，但我们尚不清楚它是否为全部教育水平e的最高盈利呢？这就要看当$e\geq e^*(H)$时的情况，请注意，当$e\geq e^*(H)$时，(13.21)式告诉我们，$w(e)=y(H,e)$，由此若低能力工人发送的教育水平信号大于等于$e^*(H)$时，其相应的盈利函数应为

$$y(H,e)-C(L,e)\quad (e\geq e^*(H)) \tag{13.22}$$

不妨设y与C均可对e求偏导数，(13.22)式的偏导数在$e=e^*(H)$处满足

$$\begin{aligned}&y_e(H,e^*(H))-C_e(L,e^*(H))<y_e(H,e^*(H))\\&-C_e(H,e^*(H))=0\end{aligned} \tag{13.23}$$

(13.23)式中的不等号是由于(13.7)式的缘故,(13.23)式中最后的等号是因为在 $e^*(H)$ 处,$y(H,e^*(H))-c(H,e^*(H))$ 为极大值的缘故,正因为 $y(H,e)-C(H,e)$ 在 $e^*(H)$ 处取得最大值,由此,有

$$y_e(H,e)-C_e(H,e)\leqslant 0 \quad 当\ e\geqslant e^*(H) \qquad (13.24)$$

因而

$$y_e(H,e)-C_e(L,e)<0 \quad 当\ e\geqslant e^*(H) \qquad (13.25)$$

这表明,当 $e\geqslant e^*(H)$ 时,$y(H,e)-C(L,e)$ 是递减函数,也就是说,低能力工人在 $e\geqslant e^*(H)$ 的范围内其盈利函数在 $e=e^*(H)$ 时取得极大值 $w^*(H)-C[L,e^*(H)]=y(H,e^*(H))-C(L,e^*(H))$。注意到我们考虑的是无嫉妒(也就无"欺骗")情况,在那儿,$w^*(L)-C[L,e^*(L)]>w^*(H)-C[L,e^*(H)]$,因此,对于一切 e 的选择,低能力工人的最优反应是 $e^*(L)$。

结论:在无嫉妒情况下,工人选取教育水平 $[e(L)=e^*(L),e(H)=e^*(H)]$、公司的信念为 $\mu(H|e)=1$(当 $e\geqslant e^*(H)$),公司的纯策略为(13.21)式,构成博弈的分离完美 Bayes 均衡。

3. 嫉妒情况下的分离均衡

前面指出,存在嫉妒情况是更有实际意义的。现在,高能力工人不能简单地通过选择在完全信息下选取的教育 $e^*(H)$ 来获得高工资 $w(e)=y(H,e)$。取而代之,为发送有关自己能力的信号,高能力工人必须选择 $e_s>e^*(H)$。这是因为低能力工人将以 $e^*(H)$ 与 e_s 之间的任何 e 值作为发送的教育水平信号,只要这样去做可以哄骗公司相信他具有高能力,他一定会发出较高信号 e。用一句通俗的话来形容,因为社会上出现低能力工人以较高学历争取到高工资的可能,高能力的工人只能以更高的学历作为信号使公司认识到他的才华。

自然的分离完美 Bayes 均衡在形式上包含了工人的策略 $[e(L)=e^*(L),e(H)=e_s]$,均衡信念 $\mu(H|e^*(L))=0$,$\mu(H|e_s)$

$=1$ 以及公司的均衡工资 $w[e^*(L)]=w^*(L)$，$w(e_s)=y(H,e_s)$。

支持上述均衡说法的公司在非均衡路径信息集上的信念是

$$\mu(H|e)=\begin{cases}0 & e<e_s\\ 1 & e\geqslant e_s\end{cases} \tag{13.26}$$

即，当 $e\geqslant e_s$ 时认为工人具有高能力，否则认为工人仅具有低能力。于是，公司的相应策略是

$$w(e)=\begin{cases}y(L,e) & e<e_s\\ y(H,e) & e\geqslant e_s\end{cases} \tag{13.27}$$

给定这个工资函数之后，低能力工人对此有两个最优反应：老老实实地选取 $e^*(L)$ 已挣得 $w^*(L)$；选取 e_s 以获工资 $y(H,e_s)$。我们假定，选择适当的 e_s，使得低能力工人对这两种反应表示无所谓，从而他的选取将有利于 $e^*(L)$（实际情况的一个解释：花了很大代价，低能力工人以高学历获得高工资，但总的盈利并不增加，倒不如老老实实地以 $e^*(L)$ 获得 $w^*(L)$ 为好）。事实上，我们只要这样去求 e_s，使得

$$y(H,e_s)-C(L,e_s)=y(L,e^*(L))-C(L,e^*(L)) \tag{13.28}$$

(13.28)式在一般情况下有解，因为工人的成本函数随 e 增长的速度将高于 y 的增长速度，因此，常能找到一个 e_s 使(13.28)式成立。现在来看高能力工人是否会偏离分离策略。由于 $e_s>e^*(H)$，而根据(13.24)式的分析，当 $e\geqslant e^*(H)$ 时，高能力工人的盈利呈下降趋势，故在这里，至少 $e>e_s$ 不是一个好的选择。那么当 $e<e_s$ 时的 e 是否会是高能力工人愿意选择的教育呢？当 $e<e_s$ 时，由(13.27)式，高能力工人也只能被公司认为是低能力而挣工资 $y(L,e)$，那么此时是否会有

$$y(L,e)-C(H,e)\leqslant y(H,e_s)-C(H,e_s) \tag{13.29}$$

成为问题的关键。回忆 e_s 的选择满足(13.28)式，即

$$y(L,e^*(L))-C(L,e^*(L))=y(H,e_s)-C(L,e_s)$$

由于当 $e<e_s$ 时，低能力工人最优选择为 $e^*(L)$，因此当 $e<e_s$ 时，

应有

$$y(L,e)-C(L,e)\leqslant y(H,e_s)-C(L,e_s) \qquad (13.30)$$

将(13.30)式改写为

$$y(L,e)\leqslant y(H,e_s)-\{C(L,e_s)-C(L,e)\} \qquad (13.31)$$

注意到教育从 e 改善到 e_s，低能力工人将比高能力工人花费更多，(即(13.7)式所蕴含的意思)，因此

$$C(L,e_s)-C(L,e)\geqslant C(H,e_s)-C(H,e) \qquad (13.32)$$

利用(13.32)式代入(13.31)式得

$$y(L,e)\leqslant y(H,e_s)-\{C(H,e_s)-C(H,e)\} \quad 当\ e<e_s \qquad (13.33)$$

(13.33)式就是(13.29)式，这表明当 $e<e_s$ 时的所有的 e 都不是高能力工人的好选择。

综上所述，我们实际上验证了所说策略与信念是存在嫉妒情况下的分离 Bayes 均衡。

4. 混杂均衡

前面提到过完美 Bayes 均衡有共用、分离与混杂之分。本小段将就劳务市场信号博弈首先对混杂均衡进行讨论。在工人分为高低能力两种类型的情况下，设高能力工人选取教育 e_h，低能力工人在选取 e_h(概率为 π)和选取 e_L(概率为 $1-\pi$)之间随机化，如前所述，低能力工人在发送信号中采用了混合策略。因此现在我们有必要将 $SR(3)$放宽到混合策略的范围，这里，$SR(3)$决定了在观察到 e_h 或者 e_l 之后公司的信念：利用 Bayes 法则得

$$\mu(H|e_h)=\frac{q}{q+(1-q)\pi} \qquad (13.34)$$

其中 q 为"自然"选取高能力出现的概率，(13.34)式表明，由于低能力工人以 π 概率选取教育 e_h，因此观察到 e_h 的概率应为 $q+(1-q)\pi$，而观察到 e_h 后公司相信工人具有高能力的概率可以利用 Bayes 法则给出(13.34)式。显然，在观察到 e_l 之后公司相信该工人具有高能力的概率为 0：$\mu(H|e_L)=0$。对(13.34)可进一步

阐述，首先，由于高能力工人总是选取 e_h，而低能力工人以概率 π 选取 $e_h(0<\pi<1)$，因此在观察到 e_h 之后公司判断工人具有高能力的可能性将增大，(13.34)式也表明 $\mu(H|e_h)>q$；其次，当 $\pi\to 0$ 时，低能力工人趋向于几乎不再与高能力工人共用一个信号 e_h，因此在观察到 e_h 之后判断工人具有高能力的可能性趋向于 1，即 $\mu(H|e_h)\to 1$（当 $\pi\to 0$ 时）；第三，当 π 趋于 1 时，低能力工人几乎总是与高能力工人共用一个信号，因此在观察到 e_h 之后，公司对工人是否具有高能力仍如"老天爷"所安排的那样去猜测，也就是说，$\mu(H|e_h)$ 趋于先验信念 q。

当低能力工人与高能力工人分离而选择 e_L 时，信念 $\mu(H|e_L)=0$ 意指他将获得工资 $w(e_L)=y(L,e_L)$，因为公司在"断定"他属于能力较低者之后，就会按照低能力的工资标准并参照他的学历付给报酬。我们可以得出的结论是，e_L 必须等于 $e^*(L)$——低能力工人在完全信息下的教育选择。为什么 e_L 必须等于 $e^*(L)$ 呢？不妨假设 $e_L\neq e^*(L)$，那么此时低能力工人的"分离"盈利等于 $y(L,e_L)-C(L,e_L)$，与 $e^*(L)$ 时的盈利 $y(L,e^*(L))-C(L,e^*(L))$ 相比较，根据 $e^*(L)$ 的定义，无可非议地成立

$$y(L,e^*(L))-C(L,e^*(L))>y(L,e_L)-C(L,e_L)$$

因此，低能力工人的"分离"信号必定是 $e^*(L)$，才能使自己获得最大可能盈利。

对于那些愿意在 $e^*(L)$ 与 e_h 两者之间随机选取的低能力工人，根据混合策略的定义，该策略必须使低能力工人在发送 e_h 还是发送 $e^*(L)$ 这两个信号之间表示出无所谓的态度，或者说，这两种教育选择导致的盈利一样多：

$$w^*(L)-C[L,e^*(L)]=w(e_h)-C(L,e_h) \qquad (13.35)$$

然而，由于 $w(e_h)=w_h$ 是公司支付的均衡工资，因此

$$w_h = \frac{q}{q+(1-q)\pi} \cdot y(H, e_h) + \frac{(1-q)\pi}{q+(1-q)\pi} \cdot y(L, e_h) \tag{13.36}$$

显然，$w(e_h)=w_h$ 应当既满足(13.35)式又满足(13.36)式。倘若给定一个 e_h 后，由(13.35)式解得的 $w_h<y(H,e_h)$，那么这个 w_h 与(13.36)式没有任何矛盾，我们完全可以根据这个给定的 e_h 解得 w_h，再解得 π。于是我们就构成了一个混杂均衡，在这个混杂均衡中，低能力工人在 $e^*(L)$ 与 e_h 两者之间随机化选取，且取 e_h 的概率为 π。假如对于 e_h 并由(13.35)式解得的 w_h 满足条件：$w_h>y(H,e_h)$，那么不无遗憾地宣布，不存在包含该 e_h 的混杂均衡。

§13.3 合作投资与资本结构

某企业家已经开办了一家公司，但需要一笔额外资金以投资一个颇有吸引力的新项目。能为人们所观察到的是公司的总计获益，但人们无法将新项目的效用与原公司的效用分开，而企业家却拥有关于原公司的盈利的私人信息。假如该企业家向一个潜在的和可能的投资者提出以公司股本换取必要的资金，在什么情况下新项目将被开发，以及投资于新项目的股本应为多少?这就是本节所要讲述的合作投资与资本结构(corporate investment and capital structure)问题。

我们可以对此建立一个信号博弈模型。原公司盈利 π 分为高(H)、低(L)两种类型，其中 $H>L>0$。为获得"新项目是吸引人的"这一想法，假如需要投资额为 I，盈利将是 R，潜在投资者另一个回报率是 r，并且 $R>I(1+r)$。博弈展开如下：

(1)"自然"确定原来公司的获益，$P(\pi=L)=p$；

(2)企业家了解 π 的知识，然后向潜在投资者提出投资于新项目的股份 s，$0\leqslant s\leqslant 1$；

(3)投资者(接收信号者)观察到 s(不是观察到 π)之后,决定要么接受提议要么拒绝提议;

(4)投资者若拒绝开价,那么投资者的效用为 $I(1+r)$;企业家的效用是 π。如果投资者接受 s,那么投资者的效用是 $s(\pi+R)$,企业家的效用为 $(1-s)(\pi+R)$。

现在,假设投资者在收到企业家的开价 s 之后,认为 $\pi=L$ 的概率为 q,即 $\mu(\pi=L|s)=q$,在这个信念下,投资者愿意投资的期望效用为 $s[qL+(1-q)H+R]$,显然,当且仅当

$$s[qL+(1-q)H+R]\geqslant I(1+r) \tag{13.37}$$

投资者愿意接受开价 s,对于企业家来说,倘若原来公司的收益是 π,接受投资者的资金 I 以后新项目的盈利为 R,显然,当且仅当

$$s\leqslant R/(\pi+R) \tag{13.38}$$

时,企业家愿意以 s 股份换取投资者的资金。

考虑在这个信号博弈中是否存在共用完美 Bayes 均衡。假如存在共用均衡,那么企业家无论 $\pi=H$ 或者 $\pi=L$,都将会向投资者提出相同的开价 s。毫无疑问,投资者在观察到 s 之后,对原来公司收益为高还是低的信念 $\mu(L|s)=q$ 应该等同于"自然"选取 $\pi=L$ 的先验概率 p,即 $q=p$。由于投资者认为企业家愿以 s 换取 I 的充要条件是(13.38)式,而 $\pi=L$ 比 $\pi=H$ 更容易使(13.38)式成立,或者说,当 $\pi=H$ 时,(13.38)式的成立相对更困难一些。一旦给定 $q=p$,综合(13.37)式与(13.38)式可知,仅当

$$\frac{I(1+r)}{pL+(1-p)H+R}\leqslant\frac{R}{H+R} \tag{13.39}$$

时共用完美 Bayes 均衡的确存在。当 $p\to 0$ 时,(13.39)式几乎成为 $R>I(1+r)$,这是我们在建立模型最初所假设的基本条件,就是说,当 p 充分地接近于 0 时,(13.39)式当然成立。故当 p 几乎为 0 时,博弈存在共用完美 Bayes 均衡。但是,当 p 充分接近于 1 时,仅

当

$$R-I(1+r)\geqslant\frac{I(1+r)H}{R}-L \tag{13.40}$$

(13.39)式才成立。读者若要验证这一点,仅需变化(13.40)式为

$$R+L\geqslant I(1+r)(1+\frac{H}{R})=\frac{I(1+r)(R+H)}{R} \tag{13.41}$$

再转化为

$$\frac{I(1+r)}{R+L}\leqslant\frac{R}{R+H} \tag{13.42}$$

当 p 充分接近于 1 时,(13.42)式左边分母 $R+L\approx R+pL+(1-p)H$。这就蕴含着(13.39)式。当原公司几乎是低效益的话,那么共用均衡指出,从投资新项目上获得的利益必须比投资者的成本与期望得到的回报一起的总数要多一定的程度。而如果投资者相信 $\pi=H$,即相信原公司是高效益的话,那么他愿意接受较少的开价股份 $s\geqslant I(1+r)/(H+R)$。在共用均衡中,以投资 I 而要求得到较大的均衡股份对于高效益公司来说过于昂贵,也许这样昂贵的索取使高效益公司宁可放弃新项目。

倘若(13.39)式不成立,则不存在共用均衡。那么是否存在分离完美 Bayes 均衡呢? 企业家的分离策略可写作 s_L 与 s_H,分别对应于低效益类型与高效益类型公司开出的股份数。如前面所分析的那样,$s_L>s_H$,因为高效益公司不愿花费昂贵代价引进新项目。容易分别计算得到 $\mu(L|s_L)=\mu(L|s_H)=1$。在给定信念 $\mu(L|s_L)=1$ 下,不难知道,低效益类型公司开出 $s_L=I(1+r)/(L+R)$,并为投资者所接受。不过此时投资无效率地低,因为投资者获得盈利 $I(1+r)$,这是他不参与该项投资就能得到的收益。在信念 $\mu(L|s_H)=1$ 下,似乎也可以使 s_H 为 $I(1+r)/(H+R)$,但是由于前提条件是:(13.39)式不成立,因此 $I(1+r)$ 与 R 几乎无差异。对于高效益公司来说,只有令 $s_H<I(1+r)/(H+R)$,才会感到可以降低昂贵的成本,否则它几乎没有挣更多的钱而多少承担一定风险,从

而不愿引进新项目，对于这样的 s_h 投资者肯定不会接受，因为他的收益还不如他不投资这个项目。这个均衡阐述了如下意义：发送者的可行信号集是无效的：高效益类型无法显示自己的特色，对高效益公司有吸引力的投资项目对低效益公司更具吸引力。

现在考虑企业家向投资者提供债务的可能性，假使投资者接受债务契约 D，如果企业家不宣告破产，那么投资者收益是 D，企业家的收益是 $(\pi+R-D)$；如果企业家宣布破产，那么投资者的收益是 $(\pi+R)$，而企业家则一无所有。考虑 $L>0$，即低效益类型的收益仍为正。此时总存在一个共用完美 Bayes 均衡，不管是高类型还是低类型均开价 $D=I(1+r)$，投资者接受这个开价（这个均衡很容易核实，读者不妨作为练习）。倘若 L 充分地负，使得 $L+R<I(1+r)$，即达到资不抵债的地步，那么投资者不会接受 D。

如果 L 与 H 不是表示确定的而是表示预期的盈利，讨论可以同样地进行。令类型 π 表示：原来公司以 $\frac{1}{2}$ 概率获益 $(\pi+K)$，以 $\frac{1}{2}$ 概率获益 $(\pi-K)$（注意：此时 $\pi=L,H$ 不是确定事件）。假如 $L-K+R<I(1+r)$，表示存在 $\frac{1}{2}$ 可能性使得低效益公司不能偿还债务，投资者当然不会接受契约。

§13.4 廉价交谈博弈

廉价交谈博弈（Cheap-talk Games）也属于信号博弈，不过又有它自己的特点，因为在廉价交谈博弈中发送者的信号几乎不涉及成本问题，它仅仅是交谈，而且这种交谈可以没有约束力。在某些信号博弈中，交谈不能作为信号，因为它不提供任何真实的信息。例如，在劳务市场模型中，面谈虽然是重要的环节，但是工人宣布“我相当能干”不足以使公司相信他是高能力工人。但是在许多其他场合，不花费成本的廉价交谈可以提供信息。尤其在较多经济

利益关系的应用中，例如由银行行长向记者发表有关"近期银行会降息纯属谣传"的讲话，实际上向客户们传递了一个虽不太精确但对客户将要作出的决策有参考价值的信息。

要使廉价交谈能具有传递信息的价值或可能，首先一个必要条件是，不同的发送者类型使接收者在行动上有不同的选择。这一点在 Spence 的劳力市场模型中就并非如此，因为无论发送信号者是高能力或是低能力，他们都喜欢高工资，如果一个廉价交谈信号（称自己能力高）可以获得高工资，而另一个廉价交谈信号（称自己能力低）导致低工资，那么每个类型的工人都因为喜欢高工资而宣称自己能力提高。这样会对接收者产生误导。因此在劳务市场模型中，不可能存在一个用廉价交谈来影响工资的均衡。上面所提到的银行行长发表政策声明则使"廉价交谈"传递了重要的信息，不同的政策声明无疑对客户的存贷款策略产生重要影响。

第二个必要条件则是接收者宁愿依赖于发送者的类型选择不同的行动。如果接收者关于行动的选择独立于发送者的类型的活，那么无论是信号博弈也好，还是廉价交谈博弈也好，它们都没有什么用处。因此第二个必要条件其实是显而易见的。

廉价交谈具有信息价值的第三个必要条件是，接收者关于行动的选择完全地不与发送者类型对立。假设对于低类型发送者，接收者喜欢选择低行动，而对于高类型发送者，接收者愿意采取高行动，又如果低类型发送者也愿意低行动，高类型发送者喜欢高行动，那么就发生了交流（或信息传递）。否则，如果发送者有对立的喜好，那么不能交流，因为会发生发送者误导接收者的事件。

现在我们介绍最简单的廉价交谈博弈的时序，它实际上与简单的信号博弈的时序相同，只不过在收益盈利方面有所不同。

(1)"自然"为发送者从可行类型集 $T=\{t_1,\cdots,t_I\}$ 中按照概率分布 $p(t_i)(i=1,2,\cdots,I)$ 抽取类型 t_i，其中对每一个 i 成立 $p(t_i)>0$，且 $p(t_1)+\cdots+p(t_I)=1$；

(2)发送者观察到类型 t_i 后，从可行信号集 $M=\{m_1,\cdots,m_J\}$ 中选取一个信号 m_j；

(3)接收者观察到 m_j(但不是 t_i)后，从可行性行动集 $A=\{a_1,\cdots,a_K\}$ 中选取行动 a_k；

(4)盈利函数分别记作 $U_S(t_i,a_k)$，$U_R(t_i,a_k)$。

在(4)中，信号对于发送者与接收者的效用没有直接作用，这正是廉价交谈博弈的关键性质。信号可以算一回事的唯一途径是通过其信息内含：通过改变接收者关于发送者类型的信念，信号可以改变接收者的行动，从而间接地影响局中人双方的效用或盈利。由于相同的信息可以用不同的语句传递，不同的信号空间可以达到同样的结果。因此廉价交谈的精神是什么事都可以说，但是把这一点形式化就要求 M 是一个非常大的集合。要做到这一点，在处理上有困难，一个取代的办法是，假定 M 足够丰富到可以说"需要说的话"，在模型的处理上，无非是令 $M=T$。在我们目前正在讨论的廉价交谈博弈中，"$M=T$"等价于允许交谈任何事情。如果涉及到今后将讨论的关于完美 Bayes 均衡的精炼，那么必须重新考虑假设。

既然最简单地廉价交谈与最简单的信号博弈有着相同的顺序，不难设想它们的完美 Bayes 均衡的定义也是相同的。在廉价交谈博弈中的纯策略完美 Bayes 均衡由一对策略$\{m^*(t_i),a^*(m_j)\}$和一个信念 $\mu(t_i|m_j)$ 组成，而且他们满足条件 $SR(1)$、$SR(2S)$、$SR(2R)$与 $SR(3)$，只不过在 $SR(2S)$与 $SR(2R)$中信号博弈的效用函数 $U_S(t_i,m_j,a_k)$、$U_R(t_i,m_j,a_k)$在这里应当换作 $U_S(t_i,a_k)$，$U_R(t_i,a_k)$。正因为效用函数中没有信号的直接影响，直观上可以想象廉价交谈博弈一定会存在共用均衡(只要均衡存在的话)，这是廉价交谈与信号博弈在结论上的重要差异之一。因为从前面知道，在信号博弈中未必存在共用均衡，常常需要满足一些条件才有可能。因为信号对发送者的盈利没有直接影响，如果接收者忽略所有

信号的话，那么共用策略是发送者的最优反应。又因为信号对接收者的盈利没有直接影响，如果发送者选取了共用策略，那么对于接收者来说，一个最佳反应是忽略所有信号。若以 a^* 表示接收者在共用均衡中最优行动，那么 a^* 因当时下述公式之解：

$$\max_{a_k \in A} \sum_{t_i \in T_i} p(t_i) U_R(t_i, a_k) \tag{13.43}$$

于是，在共用完美 Bayes 均衡中，发送者选取任何共用策略，接收者在观察到所有信号(包括在均衡路径上与非均衡路径上的)之后，坚持先验信念 $p(t_i)$ 并不管信号如何而采取 a^*。既然在廉价交谈博弈中，总是存在这样的共同完美 Bayes 均衡，自然而然地会提出一个有趣问题，除了这种总归存在的共用均衡外，还存在其他的非共用均衡吗？本节接着讨论的两个问题分别阐述了廉价交谈博弈中的分离与部分共用均衡。

取最简单的例子，发送者只有两个类型，$T=\{t_L, t_H\}$，先验概率为 $p(t_L)=p$，接收者的行动空间中也仅为两个元素，$A=\{a_L, a_H\}$。其相应各个结局的盈利或效用可以用效用矩阵形式表示如图 13.6 所示。

		发送者 t_L	发送者 t_H
接收者	a_L	1, x	0, y
	a_H	0, z	1, w

图 13.6

图 13.6 看起来只不过是一个普通的效用矩阵，但在这里却隐含着一层重要的意思：每个决局 (t_i, a_k) 的盈利与信号 m_j 没有直接关系。但是我们不要误以为它是一个正则型博弈，因为两个局中人之间还存在着“廉价交谈”式的信息传递。图 13.6 告诉我们，当发送者的类型为低 (t_L) 时，接收者若取低 (a_L) 行动则得 1，若取高行动 a_H 则得 0，因此接收者喜欢采取低行动，同样我们发现当发送

者是高类型(t_H)时,接收者将乐于采取高行动。因此图 13.6 的效用函数满足廉价交谈可传递信息的必要条件之一。现在来看第一个必要条件,假如两种发送者的类型对行动有相同的偏爱,例如 $x>z$,$y>w$,即两个类型 t_L 与 t_H 都喜欢接收者采取 a_L 而不喜欢 a_H。于是,这两种类型都将愿意接收者相信 $t=t_L$,从而使接收者的选择正是他们所喜欢的 a_L,在这种情况下,接收者不能相信这种要求。再从发送者的角度考虑问题,假如 $z>x$,$y>w$ 同时成立,表明低类型喜欢高行动而高类型喜欢低行动,于是出现了对立的情况,t_L 类型发送者将希望接收者相信 $t=t_H$ 以使接收者恰好采取 a_H,t_H 类型发送者将希望接收者相信 $t=t_L$ 以使接收者采取 a_L,其结果将使得接收者对什么都不相信。因此在这个博弈中,为使廉价交谈真正地传递信息,看来必须有 $x\geqslant z$,$w\geqslant y$。我们得到了这个廉价交谈博弈的分离完美 Bayes 策略:发送者的策略是$\{m(t_L)=t_L,m(t_H)=t_H\}$,接收者的信念是 $\mu(t_L|t_L)=\mu(t_H|t_H)=1$,接收者的策略是$\{a(t_L)=a_L,a(t_H)=a_H\}$。由于这些策略与信念构成了均衡,每一个发送者类型宁可讲真话,导致接收者采取自己喜欢的行动,而不愿说谎,以导致自己不喜欢的行动出现。

结论是,当且仅当 $x\geqslant z$,$w\geqslant y$ 时,博弈存在分离完美 Bayes 均衡。

接下来考虑的廉价交谈博弈是 Crawford-Sobel 模型的特殊情况,在那里,类型、信号与行动空间具有连续统势:发送者的类型均匀地分布于 $T=[0,1]$上,令 $p(t)$表示 t 处的概率密度,则有 $p(t)=1$;信号空间等于类型空间,即 $M=T$;行动空间也为 $A=[0,1]$,接收者的效用函数为 $U_R(t,a)=-(a-t)^2$,发送者的效用函数是 $U_s(t,a)=-[a-(t+b)]^2$。

从效用函数不难看出,如果发送者的类型是 t,那么接收者的最优行动应当是 $a=t$,此时接收者获得最大盈利 0,但是对于发送者来说,他最愿意接收者采取行动 $a=t+b$,使得他可以获得最大

盈利 0。至少，较高类型的发送者喜欢较高的行动，而且可以看到，接收者与发送者喜好或选择不是对立的，尽管当 $b>0$ 时，他们的喜好有一定差异，但是 b 越小，他们之间的"兴趣爱好"也越趋于一致。

Crawford 与 Sobel 证明了这个模型以及更广的一类有关模型的所有完美 Bayes 均衡等价于下述形式的部分共用均衡：将类型空间划分为 n 个区间 $[0,x_1),[x_1,x_2),\cdots,[x_{n-1},1]$；在每一个给定的小区间内的所有类型发送相同的信号，不同区间的类型则发送不同的信号。这里，每个小区间内所有类型的发送者采取共有策略——选择相同的信号，"部分共用"的名称也就因此而生。如果 $n=1$，那么就是"全部"共用，我们已经指出，在廉价交谈博弈中，共用均衡总是存在。现在考虑 $n=2$ 的情况，$[0,1]$ 划分为 $[0,x_1)$ 与 $[x_1,1]$ 两部分（其中 $0<x_1<1$）。此时的部分共用均衡是什么样，需要什么样的条件呢？

不妨设 $[0,x_1)$ 中所有的类型发送信号 m_1，$[x_1,1]$ 中的所有类型发送信号 m_2。在均衡中，当接收者观察到了 m_1 之后，相信发送者的类型均匀地分布在 $[0,x_1)$ 上，由 $U_R=-(a-t)^2$ 容易得出，接收者的最优行动是 $a=x_1/2$。类似地，当接收者观察到信号 m_2 之后，他相信发送者的类型均匀地分布在 $[x_1,1]$ 上，因此，接收者的最优行动是 $a=(x_1+1)/2$。两个最优行动都在各自区间的中点。再从发送者的角度考虑，要想使 $[0,x_1)$ 中所有类型都愿意发出信号 m_1，必须是这些类型都喜欢行动 $x_1/2$ 而不喜欢 $(x_1+1)/2$，同样，要使 $[x_1,1]$ 中所有类型喜欢发送信号 m_2，必须是这些类型喜欢行动 $(x_1+1)/2$ 而不喜欢 $x_1/2$。

注意到发送者的效用函数是 $U_s=-[a-(t-b)]^2$，他最喜欢的行动是 $(t+b)$，效用函数是个顶点在 $(t+b)$ 开口向下的抛物线。用图 13.7 简单且直观地解释这个问题。

给定 $n=2$ 的区间划分 $[0,x_1)$ 与 $[x_1,1]$ 后，对于任意类型 t 的

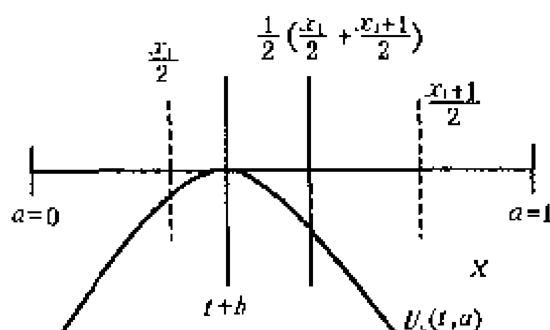


图 13.7

发送者来说，他最喜欢的行动$(t+b)$有三种可能：位于$\frac{x_1}{2}$与$\frac{x_1+1}{2}$两点之间的中点$\frac{1}{2}(\frac{x_1}{2}+\frac{x_1+1}{2})$上；位于该中点的左边或右边。由图 13.7 可以看出，如果$t+b<\frac{1}{2}(\frac{x_1}{2}+\frac{x_1+1}{2})$，类型$t$发送者喜欢$\frac{x_1}{2}$而不喜欢$\frac{x_1+1}{2}$，此时，$t$应属于$[0,x_1)$，反过来可以知道，如果$t+b>\frac{1}{2}(\frac{x_1}{2}+\frac{x_1+1}{2})$，类型$t$发送者喜欢$\frac{x_1+1}{2}$而不喜欢$\frac{x_1}{2}$，这种情况，从均衡角度来看，应有$t\in[x_1,1]$。因此，欲存在均衡，那么两个小区间的分界$x_1$点应当对行动$\frac{x_1+1}{2}$与$\frac{x_1}{2}$喜欢程度无所谓，也就是说，必须有

$$x_1+b=\frac{1}{2}(\frac{x_1}{2}+\frac{x_1+1}{2}) \tag{13.44}$$

或者

$$x_1=\frac{1}{2}-2b \tag{13.45}$$

总结一句话，为使二步均衡存在，x_1必须是这样的类型t，使得他的最优行动$t-b$恰好是$\frac{1}{2}(\frac{x_1}{2}+\frac{x_1+1}{2})$。因为$T=[0,1]$，故$x_1$必须为正，由(13.45)式，仅当$b<1/4$时才存在二步均衡。

看来，在$b<1/4$时有可能存在“二步”部分共用均衡，发送者

的策略是，当 $t\in[0,x_1)$ 时，通过廉价交谈发出自己在 $[0,x_1)$ 的信号，接收者在观察到这个信号（例如听到对方有关自己位于 $[0,x_1)$ 区间的声明）之后认为发送者的确属于 $[0,x_1)$ 类型，从而采取行动 $\frac{x_1}{2}$；当 $t\in[x_1,1]$ 时，发送者可以通过"声明"等廉价交谈手段发出自己在 $[x_1,1]$ 的信号，接收者观察到这个信号后认为发送者确属此类型，采取行动 $\frac{x_1+1}{2}$。

上述二步均衡从理论上认为是可能的，我们沿用一个类似的游戏试图解释这个可能的部分公用均衡。设有甲乙两人，甲从计算机上利用 $[0,1]$ 上均匀分布抽得随机数 t，乙猜测它是 a，甲的效用为 $U_s=-[a-(t+b)]^2$（对某正数 b），乙的效用为 $U_R=-(a-t)^2$。该游戏有平凡的共用均衡解，乙猜测 $a=1/2$。如果 $n=2$，当 $b<1/4$ 时，甲可以告诉乙随机抽得的 t 是在 $[0,1/2)$ 还是在 $[1/2,1]$ 内，二步共用均衡是乙若听到甲说 t 在 $[0,1/2)$ 内，则猜 $a=1/4$，否则猜 $a=3/4$。读者不难发现，在游戏进行之前确定 $b=0$ 的话，在规定 $n=2$ 的前提下，甲取 $x_1=1/2$ 为最优。但 $0<b<1/4$ 时，依前所言，甲取 $x_1=1/2-2b$。

现在我们应该进一步完善两步完美 Bayes 均衡，自然而然地寻找不在均衡路径上的信号事件。Crawford 与 Sobel 指定了一个发送者的（混合）策略，使得这样的信号并不存在：对于所有小于 x_1 的类型，按照 $[0,x_1)$ 上均匀分布规律随机地选取一个数作为廉价交谈发送的信号，例如若随机抽得为 $x_1/3$，则发送者可宣布"类型为 $x_1/3$"；对于所有 $t\geqslant x_1$ 的类型，则按照 $[x_1,1]$ 上均匀分布规律随机地选取一数作为信号通过廉价交谈发送。因为我们假定信号空间 M 就是类型空间 T，因此实际上不可能存在信号不在均衡路径上。这样 $SR(3)$ 确定了接收者在观察到所有可能信号后的信念：如果观察到来自 $[0,x_1)$ 中的任何信号，那么接收者的信念为"t

在$[0,x_1)$均匀地分布”，如果观察到来自$[x_1,1]$的任何信号，接收者的信念为“t在$[x_1,1]$上均匀地分布”。与Crawford与Sobel的办法相对立地，可以指定发送者的一个纯策略而使接收者选取一个适当的不在均衡途径上的信念。例如，令发送者的策略为，$t<x_1$的所有类型发送信号0，$t\geqslant x_1$的所有类型发送信号x_1，这样除了0与x_1之外，其余所有类型都不在均衡途径上。此时接收者在观察到信号后绝不会认为发送者类型只是0，他的非均衡路径信念为“t均匀地分布在$[0,x_1)$上”，同样，在观察到x_1之后接收者的非均衡路径信念为“t均匀地分布在$[x_1,1]$上”。

前面列举游戏例子分别可以描述为上述两种策略剖面与后验信念。例如，当甲抽得的随机数$t\in[0,1/2)$时，可以发出信号“t在$[0,1/2)$上均匀地分布”，也可采用纯策略发送信号“0”。乙在观察到这两种结果之后的信念均为“t在$[0,1/2)$均匀地分布”。

继续游戏下去，设想n不再是2，而是大于2的正整数，于是我们将$[0,1]$区间分为$[0,x_1)$，$[x_1,x_2)$，…，$[x_{n-1},1]$共n个区间，如果甲抽取某数后，通过廉价交谈发送出更精确信息的信号，乙树立了甲究竟在哪个小区间的信念以后所采取的行动——以该小区间的中点作为对t的猜测——想必会使甲乙两人取得更有效的盈利。这种想法有点像定积分引进时的划分小区间计算曲线下近似的矩形面积，然后取极限得到曲边矩形面积的精确值。读者也许会想到，如果将$[0,1]$区间n等分，当n越来越大。或者说是趋于无穷时，部分共用均衡似乎也就渐渐趋向于甲真实地汇报自己所抽取到的数t，由于信息的百分之百的真实性，乙相当正确地猜测到甲的t，也就使两人的盈利相当理想——可惜这仅仅是如意算盘。由于正数b的存在，甲讲真话未必构成均衡，因为乙采用行动$a=t$使自己的盈利极大化，这对甲显然不利，他会偏离“讲实话”以使自己的盈利有所提高。同样的理由也会使人们想到，由于b的存在，n如果很大的话仍然可能达不到部分共用均衡，n究竟取多大使得

部分共用均衡存在且效率又高，看来与 b 的大小有关系，不妨记为 $n(b)$。下面我们将对这个问题进行研究。

为描述 n 步部分共用均衡，先回过头来考察 $n=2$，我们发现，如果 b 给定，在二步均衡中，应有 $x_1=1/2-2b$，即 $[0,x_1)$ 的区间长度为 $(1/2-2b)$，而后一区间 $[x_1,1]$ 长度为 $1-x_1=1/2+2b$。后一区间比前面区间长 $4b$。二步情况中二个区间长度不一样（随 b 而定），那么多步情况是否也应该是后面区间长于前面的区间呢？设想在多步均衡情况中，毗连的二个小区间具有相同的长度，现在发送者的类型 t^* 恰好在二个小区之间的边界上（相当于二步均衡中 x_1 的地位），由于二个区间一样长，接收者对以这二个小区间信号的最优反应分别是二个区间的中点，这二个"中点"离边界 t^* 的距离一样，由于发送者最喜欢的行动是 t^*+b，从效用函数的形状（一个向下的倒置抛物线）知道，类型 t^* 发送者一定喜欢发送出与前面区间相关的信号。其实，每两个等长度的毗连区间的"边界类型"发送者都有这种偏爱习性。为使发送者与接收者的策略剖面成为均衡，这些"边界类型"发送者必须对到底发送与上一个或下一个区间中哪一个区间有关的信号感到无所谓，于是，不管 n 等于多少，在均衡状态，每后一个区间都应比紧接它的前一区间长一些。具体分析如下：

如果 $[x_{k-1},x_k)$（第 k 个小区间）的区间长度等于 c，$x_k-x_{k-1}=c$，那么当发送者发送出该区间类型的共用策略后，接收者的最优反应是 $(x_{k-1}+x_k)/2$，这个点必定位于边界类型 x_k 所最喜欢的行动 x_k+b 以下。

$$x_k+b-\frac{1}{2}(x_{k-1}+x_k)=\frac{1}{2}(x_k-x_{k-1})+b=\frac{c}{2}+b \qquad (13.46)$$

为使边界类型 x_k 对发送与 $[x_{k-1},x_k)$ 和与 $[x_k,x_{k+1})$ 有关信号感到无所谓，那么接收者对与 $[x_k,x_{k+1})$ 有关的信号作出的最优行动必定在类型 x_k 最喜欢的行动 (x_k+b) 之上 $(\frac{c}{2}+b)$：

$$\frac{1}{2}(x_k+x_{k+1})-(x_k+b)=\frac{c}{2}+b \qquad (13.47)$$

或者

$$x_{k+1}-x_k=c+4b \qquad (13.48)$$

就是说，后面区间$[x_k, x_{k+1})$比前面区间$[x_{k-1}, x_k)$长 $4b$，由于 k 的任意性，因此在 n 步均衡中（只要 n 是允许的），每一步均比前面一步长 $4b$。设第一步（即第一个区间$[0, x_1)$）具有区间长度 d，那么第二步必定为 $d+4b$，第三步为 $d+8b$，等等，一直下去。第 n 步的终点为 1，于是所有 n 步区间总长应当等于 1：

$$d+(d+4b)+\cdots+[d+(n-1)\cdot 4b]=1 \qquad (13.49)$$

注意到 $1+2+\cdots+(n-1)=\frac{1}{2}n(n-1)$，我们有如下等式：

$$nd+2b\cdot n(n-1)=1 \qquad (13.50)$$

如果 n 满足条件 $2b\cdot n(n-1)<1$，那么(13.50)式也给予我们一个合理的 d 值。此时显然存在一个 n 步部分共用均衡。由(13.50)式可知，由于 d 必须大于 0（否则没有第一小区间），因此在给定 b 条件下且满足 $2b\cdot n(n-1)<1$ 的最大 n 值不难从下式求得：

$$2bx^2-2bx-1=0 \qquad (13.51)$$

x 的正根值应为

$$x^*=\frac{2b+\sqrt{4b^2+8b}}{4b}=\frac{1}{2}\left(1+\sqrt{1+\frac{2}{b}}\right) \qquad (13.52)$$

因此 $n^*(b)$ 应为 $\frac{1}{2}\left(1+\sqrt{1+\frac{2}{b}}\right)$ 的整数部分（若 $\frac{1}{2}\left(1+\sqrt{1+\frac{2}{b}}\right)$ 恰为整数，则需减去 1）。

在二步部分共用均衡中，我们已经指出，$b<1/4$是必须的。发送者与接收者喜欢的行动之间的差是 b，因此 b 越小，发送者与接收者的偏好选择越一致，b 越大，这二人的选择喜好差距也越大。

当 $b \geqslant 1/4$ 时

$$\frac{1}{2}[1+\sqrt{1+\frac{2}{b}}] \leqslant \frac{1}{2}[1+\sqrt{9}]=2 \qquad (13.53)$$

此时 $n^*(b)=1$。$n=1$ 时发送者与接收者之间其实不存在廉价交谈的内容，因为发送者的唯一共用策略对接收者毫无意义。事实上，当 $b \geqslant 1/4$ 时，$n^*(b)=1$ 蕴含着如果局中人的选择不太一致的话，局中人之间不可能有交流。(13.52)式也蕴含着 $n^*(b)$ 随 b 增加而减少，仅当 b 趋于 0 时 $n^*(b)$ 趋于无限。这表明当局中人的选择更接近于一致时，更多的交流可以通过廉价交谈发生。除非局中人的选择完全地一致，否则不可能发生完美的交流。

第十四章　完美 Bayes 均衡应用

§14.1　有限重复囚徒窘境的信誉效应

在第八章有关完全信息的有限重复博弈分析中，已经指出，如果阶段博弈有唯一 Nash 均衡，那么有限重复博弈就存在唯一子博弈完美 Nash 均衡：在每个历史之后的每个阶段，实施阶段博弈的 Nash 均衡。以囚徒窘境阶段博弈为例，每个局中人有合作(即抗拒交代)与背叛(即坦白)两个策略，相应的盈利矩阵重述如图 14.1 所示。

		局中人 2	
		背叛	合作
局中人 1	背叛	−8，−8	0，−15
	合作	−15，0	−1，−1

图 14.1

该阶段博弈有唯一 Nash 均衡：(背叛，背叛)，因此从理论上讲，T 阶段重复囚徒窘境的唯一子博弈完美均衡为：在每个阶段每个局中人都背叛。然而大量的实验或实践迹象表明，令警方头痛的是在有限重复囚徒窘境中"合作"经常性地发生。1982 年由 Kreps，Milgrom，Roberts 和 Wilson 所建立的信誉(reputation)模型(有时称为 KMRW 模型)对这种现象提供了很合理的解释。其实，在

第八章中，我们已经讨论过，“合作”可以发生在无限重复囚徒窘境博弈中。虽然在那里，双方局中人的盈利与机会是共同知识，这与本节所展开的情况有所不同，但是有些博弈研究工作者仍称这样的均衡为“信誉”均衡。

首先我们对信誉这个概念略作解释：在同样的阶段博弈重复进行若干次时，关心的一个问题是，局中人可能显现出取某类行动的“信誉”。为什么关心这样的问题呢？因为如果局中人总是以同一方式采取行动，那么他的对手会相应地调整自己的行动。于是研究的问题为，一个局中人在什么时候或者他是否将显现（或保持）他所希望的信誉。例如，如果中央银行总是履行它所宣布的金融货币政策，交易者将得出结论以相信中央银行在将来也会这样做吗？就是说，信誉效应将允许中央银行有效地约束自己以承担履行自己的所说的义务吗？

在有限重复囚徒窘境问题中，就存在着此类信誉问题。如果每个局中人都是“明智的”，毫无疑问地每个阶段中每个局中人都将采取“背叛”。但是倘若局中人为“明智”的概率不是 1，而存在着一定的概率 ϵ（可能比较小但却为正数）采取“针锋相对”策略，即“不管怎样，这次我取的行动是你在上次所采取的行动”，那么就存在着“合作”的可能。Kreps 等人引入了关于局中人类型的不完全信息以对局中人关心他们的信誉的可能性进行建模。在这样的模型中，每个局中人的信誉概括为其对手关于他的类型的当前信念。例如，对中央银行信守其发布的金融货币政策的信誉建模，那么对总是履行诺言的类型将赋予正先验概率。一般地，假设每一个局中人有若干不同类型，每个类型与不同种类的行动相连，并假设没有局中人的类型会作为元素直接进入任何其他局中人的盈利函数。对于有限重复囚徒窘境，其关于信誉的建模也可以将信誉同完全信息重复博弈的均衡策略等同起来。例如，以“冷酷”策略“合作直至对手背叛为止，然后一直背叛”的均衡可以解释为每个局中人都有

着关于合作的信誉，不过这种合作随着对手的第一次背叛而消失。

现在我们来观察信誉效应对有限重复囚徒窘境的影响。假设局中人关于自己的可行策略有私人信息。例如局中人1，我们假设他以概率 p 仅取针锋相对策略，而以概率 $(1-p)$ 取完全信息重复博弈中任何可取的策略，包括针锋相对在内。即局中人1有两种类型，后者称为"明智"的，前者称为"非明智"的。如果局中人1永远地偏离针锋相对策略，那么他就是我们在完全信息博弈中所熟悉的理性人。其实，针锋相对策略并非不宜采用的，相反，它是一种既简单又具有一定吸引力的策略，在某些场合也许很奏效。我们将仅取针锋相对策略的局中人类型称为非明智的，因为一个局中人假如只有一个可用策略的话，一般来说不是一个好主意。日常生活中有许多例子可以用来阐述。例如，一个棋手开局总是使用同一套路，尽管从布局来说，这不失为一种好布局之一，但是这使他的对手在重复博弈的以后阶段（即下次再下棋时）只需要研究他的一种开局就可以了。假如他在某些时候变动一下布局方式也许更好，说不定可以出奇制胜。KMRW 模型分析的主要精神在于，即使 p 非常小——即局中人2对于局中人1可能为非明智这一点只有极小怀疑——在下述意义下这种小小的不确定性却有着大的影响。Kreps 等人对序贯均衡证明了，任何一个局中人背叛的阶段数目在均衡中存在着一个上界，这个上界依赖于 p 以及依赖于阶段博弈的盈利，但是它不依赖于重复博弈中总的阶段数。根据这个结论，如果重复博弈足够地长，那么在均衡中两个局中人在相当多的阶段里存在着合作。Kreps 等人的这个结论对完美 Bayes 均衡同样适用。在 KMRW 模型中讨论的两个关键之处是：

(1)如果局中人1永远地偏离针锋相对策略，无疑局中人1是"明智的"、"理性的"就成为大家的共同知识。有限重复囚徒窘境中如果两个局中人都是明智的，那么第八章的讨论指出没有一个局中人会在此后采取合作策略，因此，明智的局中人有激情会去摹仿

针锋相对。

(2)如果对阶段博弈的盈利强加上若干假设(这一点在下面的讨论中会给出),局中人2对于针锋相对的策略的最优反应将是合作下去直到博弈的最后一个阶段。

现在我们先假设 p 足够地大,大到存在一个均衡,其中两个局中人在一个“短”的重复博弈的最后两个阶段之外的所有阶段都合作。于是我们先从两周期情况开始讨论。博弈的顺序如下:

(1)自然选择局中人1的类型。以概率 $p(0<p<1)$、局中人1为“非明智的”,他只有一个针锋相对策略可以利用;而局中人1为“明智”或“理性”的概率为 $(1-p)$,该类型可以取任何可行策略。局中人1当然知道自己属于何种类型,但是局中人2不知道对手的类型(注:我们目前讨论的情况暂时限于局中人2是明智的)。

(2)局中人1与局中人2进行第一周期囚徒窘境博弈。该阶段的行动成为下一阶段的共同知识。

(3)局中人1与局中人2进行第二次也就是最后一次囚徒窘境博弈。

(4)两个局中人在博弈中的盈利应为各自在阶段博弈中盈利的折扣和(这里暂考虑折扣因子 $\delta=1$ 的情况)。

我们提到在 KMRW 模型中对阶段博弈的盈利有所假设,因此将图14.1中盈利矩阵写成如图14.2形式。

		局中人2 背叛	局中人2 合作
局中人1	背叛	$-8,-8$	a,b
	合作	b,a	$-1,-1$

图 14.2

为使图14.2成为囚徒窘境的盈利矩阵,必须假定 $b<-8$, $a>-1$。

我们从最后一个阶段博弈开始向后倒退进行讨论，为方便起见，"合作"策略以字母 C(Cooperate)表示，背叛或告密用字母 F(Fink)表示。单独考虑最后阶段囚徒窘境时，显然局中人 2 与明智的局中人 1 都会取策略 F，这是因为 F 严优于 C 的缘故。对于明智的局中人 1 来说，他考虑到在第二阶段局中人 2 肯定取 F，因此自己没有任何理由在第一阶段取 C。根据假设，针锋相对类型的局中人 1 从合作(C)开始博弈，由于局中人 2 不知道局中人 1 究竟属于何种类型，权且假定他在第一阶段的行动为 x(x 取 C 或 F)，博弈进行情况如图 14.3 所示。

	$t=1$	$t=2$
针锋相对	C	x
明智局中人 1	F	F
局中人 2	x	F

图 14.3

通过选择 $x=C$，局中人 2 在第一阶段的期望盈利是 $-1\cdot p+(1-p)\cdot b$，而在第二阶段的期望盈利是 $p\cdot a-8(1-p)$。通过选择 $x=F$，局中人 2 在第一阶段的期望盈利是 $p\cdot a-8(1-p)$，而在第二阶段的期望盈利是 -8。由于 $\delta=1$，因此当

$$-p+(1-p)b\geqslant -8 \tag{14.1}$$

时，局中人 2 将在第一个局期采取策略行动 C。以图 14.1 为例，$b=-15$，此时若 $p\geqslant 1/2$，则(14.1)式满足，就是说，如果局中人 2 以较大把握($p\geqslant 1/2$)认为局中人 1 是仅采用针锋相对的非明智类型，那么他会在第一周期采取合作策略。在下面的讨论中，我们不妨假定(14.1)式总是成立。

现在进一步考虑三周期情况。在第一周期，针锋相对类型总是以 C 开始博弈，因此只考虑明智的局中人 1 与局中人 2 的行动选择。假如局中人 2 与明智的局中人 1 在第一周期都取合作，而且在

条件(14.1)式得到满足的情况下，显然在第二及第三周期的均衡路径应如图 14.3，于是三个周期的博弈进程将如图 14.4 所示。

	$t=1$	$t=2$	$t=3$
针锋相对	C	C	C
明智局中人 1	C	F	F
局中人 2	C	C	F

图 14.4

要使图 14.4 构成三周期博弈的均衡，我们必须探讨局中人 2 与明智的局中人 1 在第一周期取合作策略的充分条件是什么。

在图 14.4 的路径中，明智的局中人 1 的盈利应为

$$-1+a-8=a-9 \quad (14.2)$$

而局中人 2 的期望盈利应为

$$-1+[-p+(1-p)\cdot b+p\cdot a-8(1-p)] \quad (14.3)$$

由于针锋相对类型一定在第一周期取 C，因此如果明智的局中人 1 在第一周期就取 F 的话，就暴露了他的"明智"面目，这样两个局中人在第二与第三周期必然地都采取背叛(F)行动。明智的局中人 1 由于在第一周期的自我暴露，从而至多获得总盈利为

$$a-8-8=a-16 \quad (14.4)$$

它肯定小于 $a-9$，因此明智的局中人 1 对于第一周期的自我暴露不会感兴趣，他宁肯遵照图 14.4 的博弈路径且不愿主动偏离。现在的问题是，局中人 2 有没有偏离的激情呢？如果局中人 2 在第一周期就取 F，那么当局中人 1 属于针锋相对类型时肯定将在第二周期取 F，而由于局中人 2 在最后一个周期肯定取 F，故明智的局中人 1 在第二周期将取 F。在第一周期取 F 之后，局中人 2 面临的问题是在第二周期究竟取 C 还是取 F。如果局中人 2 在第二周期仍取 F，那么针锋相对将在第三局期取 F，博弈进展如图 14.5 所示。

	$t=1$	$t=2$	$t=3$
针锋相对	C	F	F
明智局中人 1	C	F	F
局中人 2	F	F	F

图 14.5

在这种情况下，局中人 2 的盈利将等于

$$a-8-8=a-16 \tag{14.5}$$

假定

$$-1-p+(1-p)b+p\cdot a-8(1-p)\geqslant a-16 \tag{14.6}$$

的话，那么局中人 2 将在第一周期不偏离 C 且使博弈进程遵照图 14.5 的形式进行，考虑到条件(14.1)式，因此局中人 2 不取如图 14.5 那样的偏离方式的充分条件可以写成：

$$-1+p\cdot(8+a)\geqslant a \tag{14.7}$$

当然我们还必须考虑局中人 2 在第一周期发生偏离而取 F，但在第二周期却又取 C 的情况，在该情况下，针锋相对在第三周期将取 C，相应的博弈进程如图 14.6 所示。

	$t=1$	$t=2$	$t=3$
针锋相对	C	F	C
明智局中人 1	C	F	F
局中人 2	F	C	F

图 14.6

局中人 2 来自如图 14.6 那样的偏离的期望盈利为

$$\begin{aligned}&p[a+b+a]+(1-p)[a+b-8]\\&=(a+b)+p\cdot a-8(1-p)\end{aligned} \tag{14.8}$$

假如

$$-1-p+(1-p)b+p\cdot a-8\cdot(1-p)$$

$$\geqslant (a+b)+p\cdot a-8(1-p) \tag{14.9}$$

那么局中人 2 的期望盈利小于"均衡"期望盈利，仍然考虑到条件(14.1)式，于是局中人 2 不取如图 14.6 形式偏离的充分条件为

$$a+b\leqslant -9 \tag{14.10}$$

综上所述，如果(14.1)式、(14.7)式与(14.10)式同时成立，那么如图 14.4 所示的博弈进程是三周期囚徒窘境的完美 Bayes 均衡的均衡路径。对于每一个给定的 $p\in(0,1)$，a 与 b 必须满足(14.1)式、(14.7)式与(14.10)式才能使博弈过程图 14.4 成立完美 Bayes 均衡的均衡路径。设想，如果 p 趋于 0，那么(14.1)式趋于 $b\geqslant -8$，而(14.7)式趋于 $a\leqslant -1$，这与囚徒窘境阶段博弈盈利矩阵的要求 $b<-8$，$a>-1$ 有悖，这表明当 p 趋于 0 时，满足三个充分条件的 a 与 b 的可能性将越来越少。

这里，我们分析的情况是"p 具有一定的正量"，在不太长的重复囚徒窘境中具有均衡合作。KMRW 则更关心于 p 值较小时重复博弈周期较长的情况。其实，对于适当大的 p 值，即使重复局期很长，仍像 $T=3$ 时那样存在着局中人合作的完美 Bayes 均衡。为简短见，如果 T - 周期重复囚徒窘境中明智的局中人 1 与局中人 2 双方一直合作直到 $T-2$ 周期，而最后两周期的博弈路径如图 14.4 所示，由此构成的完美 Bayes 均衡将被称为合作均衡(cooperative equilibrium)。$T=3$ 时的合作均衡我们已经得到，对于 $T=4,5\cdots$，乃至任何有限的 T 是否也存在这样的合作均衡呢？下面我们将证明这个结论的正确性，即如果 a,b,p 满足条件(14.1)式、(14.7)式与(14.10)式，那么对每一个 $T>3$，总存在一个合作均衡。最容易想到的办法为数学归纳法。$T=3$ 时已经知道事实为真，今假设在 $\tau=3,4,\cdots,(T-1)$ 时上述结论为真，我们试图证明在 T 周期重复囚徒窘境中也存在一个合作均衡。

为证明合作均衡的确存在，无非是从明智的局中人 1 与局中人 2 的各自角度出发，论证谁都没有激情偏离合作均衡。首先从明

智的局中人1出发。给定局中人2取合作均衡策略，即在任何$t<(T-1)$局期局中人2取合作(C)，倘若明智的局中人1在任何$t<(T-1)$周期取F，无疑"暴露"了自己的"明智"类型，于是他将在t局期获盈利a，而在以后的所有周期内接受盈利(-8)(因为双方均取F)；如果明智的局中人1不偏离合作均衡，他在t至$(T-2)$周期均获盈利(-1)，而在$(T-1)$周期获盈利a，最后一个周期获盈利(-8)。显然在$t<(T-1)$取F对明智的局中人1不利。

接下来将论证局中人2同样没有激情去偏离合作均衡策略。需要证明对于任意的$t\leqslant T-1$，局中人2不愿在t周期取F(在这之前均取C)。对图14.3及图14.4的讨论告诉我们，这里仅需假设$1\leqslant t\leqslant T-3$。

设想局中人2若在周期t取F，再在$(t+1)$周期取C，回到合作均衡策略(注意：这里我们应用了一步偏离准则，只考虑局中人2在$1\leqslant t\leqslant T-3$中任何一个阶段单独偏离的情况，有关一步偏离准则的原理，请读者回忆第八章内容)。由于局中人2在周期t时取F，必然引起针锋相对类型的局中人1在$(t+1)$周期取F，而且，就连明智类型的局中人1也在$(t+1)$局期取F，否则他若取C的话就会暴露了自己的明智类型从而引起以后一切周期的均衡结局均为(F,F)，从总盈利角度来看，显然他取C一定严劣于取F。要看到一个事实，因为两种类型的局中人1都是取C直到周期t，然后两种类型在周期$(t+1)$都取F，因此局中人2在周期$(t+2)$博弈开始之前的信念仍然是"局中人1属于针锋相对类型的概率是p"，与博弈刚开始时的信念没有发生改变。假如局中人2在周期$(t+1)$又会到合作均衡策略取C，那么从周期$(t+2)$开始的后续重复博弈将无疑地相当于τ周期$(t=T-(t+2)+1)$的有限重复囚徒窘境。归纳假设告诉我们，在进行这样的τ一周期后续重复博弈时必定存在合作均衡。现在我们来计算局中人2在t周期偏离(一步)与不偏离时分别获得的期望盈利。如果局中人2在t周期取F

而在$(t+1)$周期又取C，他从t周期开始的期望盈利为

$$a+b+[T-(t+2)+1]\text{周期后续博弈的合作均衡的期望盈利} \quad (14.11)$$

而局中人2在t周期不偏离合作均衡的期望盈利为

$$-2+[T-(t+2)-1]\text{周期后续博弈的合作均衡的期望盈利} \quad (14.12)$$

由于$a+b<-2$为显然，因此(14.11)式显然小于(14.12)式，因此我们实际上证明了局中人2不会有激情通过在$1\leqslant t\leqslant T-3$种的任何一步$t$偏离C而在$(t+1)$周期又取C这样的办法偏离合作均衡。由一步偏离准则，我们完全证明了需要得到的结论。

§14.2 非对称信息下的序贯讨价还价

在第六章中我们讨论了完全信息的讨价还价模型，本节将探讨不完全信息下的序贯讨价还价问题。我们以一个(西方国家)公司与工会之间在工资问题上的谈判为例进行介绍。为简便起见，不妨假设工作是固定的。工会成员如果不被公司雇用，其保留工资为w_r。公司的盈利以π表示，它是公司的私人信息，也就是说，只有公司知道π的真值。但是工会不知道π的真值，只可能认为π在π_L与π_H之间的任何一个值都有同样的可能性，即π服从$[\pi_L,\pi_H]$上的均匀分布。为简化分析，假设$w_r=\pi_L=0$。

设讨价还价的谈判至多持续两个周期。在第一个周期，工会开出一个工资价目，例如为w_1。假如公司接受这个开价，那么博弈宣布结束，工会盈利为w_1而公司的盈利为$\pi-w_1$。当然一般这样的工资谈判结束后将会签订数年合同，为讨论方便起见，假设合同中已经考虑了诸如折扣因子等因素。倘若公司拒绝工会提出的w_1，那么博弈进人第二个周期，在这个周期内，工会作出另一个开价w_2。如果公司接受w_2，考虑到贴现与两周期之间的生活费用等情

况，局中人盈利（以第一周期所测定的）现值为：工会 δw_2，公司 $\delta(\pi-w_2)$。如果公司拒绝工会的第二次开价，那么博弈结束，此时两者的盈利都等于 0。

在这个模型中，确定并得到一个完美 Bayes 均衡有点复杂，然而最终答案简单且直接。因此这里先通过概述该博弈模型的唯一完美 Bayes 均衡作为讨论问题的开始。

(1)工会在第一个周期的工资开价是

$$w_1^*=\frac{(2-\delta)^2}{2(4-3\delta)}\pi_H \tag{14.13}$$

(2)如果公司的盈利 π，超过

$$\pi_1^*=\frac{2w_1}{2-\delta}=\frac{2-\delta}{4-3\delta}\pi_H \tag{14.14}$$

那么公司接受 w_1^*；否则，公司拒绝 w_1^*。

(3)如果工会的第一个周期的开价遭到拒绝，工会调整关于公司盈利的信念：工会相信 π 均匀地分布在$[0,\pi_1^*]$上。

(4)工会的第二周期工资开价（在 w_1^* 被拒绝的前提下）是

$$w_2^*=\frac{\pi_1^*}{2}=\frac{2-\delta}{2(4-3\delta)}\pi_H<w_1^* \tag{14.15}$$

(5)如果公司的盈利 π 超过 w_2^*，那么公司接受工会提出的工资要求 w_2^*；否则就拒绝。

按照上述“均衡”，在每一个周期，高效益公司接受工会的开价而低效益公司则拒绝工会的开价。工会在第二周期的信念考虑了高效益公司接受第一周期工资开价的事实。在均衡中，低效益公司为了使工会确信公司是属于低效益的，不惜承受一个周期的损失，以换取工会在第二周期开出较低工资。然而，如果公司的效益相当差的话，它将发现即使工会在第二周期的开价较低也使它无法承受，从而不得不拒绝。

我们的分析先从描述局中人的策略与信念开始，从而确定完

美 Bayes 均衡。在现实生活中，公司的类型 π 以及工会的开价 w 可以有许多种可能，若用博弈树展开博弈的过程比较困难，图 14.7 所展示的博弈树仅考虑了最简单的形式：公司的类型只有两种——π_H 与 π_L；工会的工资开价也只有两个——w_L 与 w_H。

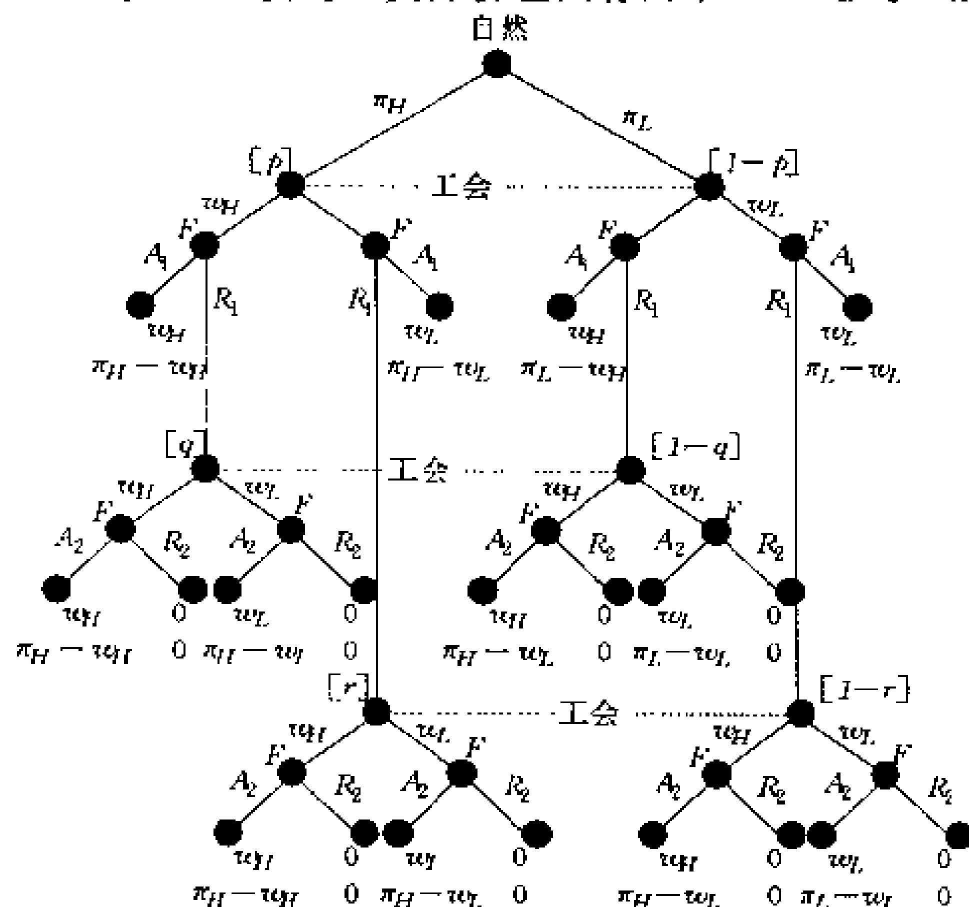


图14.7

在这个简化了的博弈中，工会在三个信息集上具有行动，因此工会的策略由三个工资开价组成：第一周期的开价 w_1 和两个第二周期的可能工资开价 w_2，即在 $w_1=w_H$ 被拒绝之后开出的 w_2 和在 $w_1=w_L$ 被拒绝之后的 w_2。这三个行动发生在三个非单一信息集上，因此在每个信息集上工会都应有关于公司类型的信念，我们分别以 $(p,1-p)$、$(q,1-q)$ 和 $(r,1-r)$ 表示，其中向量的第一个元素相应于工会认为公司为高效益的可能性。如果博弈并不像图 14.7 那样呈简单形式，那么工会的策略应该是第一周期的工资开

价 w_1 再加上第二周期的工资开价函数 $w_2(w_1)$，这个函数确定了在每一个可能的第一周期开价 w_1 被拒绝之后在第二周期提出的工资要求 w_2。这些行动当然也发生在非单一性的信息集上。但是此时情况将比图 14.7 复杂得多，因为对应于工会在第一周期的每一个可能开价 w_1 都有一个第二周期的信息集，因此存在着这些信息集的连续统，而不是图 14.7 中那样只有两个第二周期信息集。于是，倘若我们考虑的是完整的(即非图 14.7 中那样简单的)博弈，那么在第一周期内，对于每一个可能的 π 值，一定存在着工会的一个决策结，由于 π 值可能不只是两个，甚至可能是连续统的，因此工会在第一周期的信念应该是这么多结上的概率分布，不妨以 $\mu_1(\pi)$ 表示之。在第一周期开价 w_1 被拒绝之后，第二周期的信息集仍然面对那么多的 π 值，因此工会在第二周期的信念应当是相应于每一个 π 值的决策结上的概率分布，但是这个概率是有条件的——w_1 被拒绝，所以可以用 $\mu_2(\pi|w_1)$表示。

现在再来看公司的策略，无论是图 14.7 那样简化了的还是原先应考虑的完整博弈，它总是包含了两个决策：接受或拒绝。令 $A_1(w_1|\pi)=1$，表示当公司的效益类型为 π 时它在第一周期接受开价 w_1；如果他拒绝 w_1 的话，则用 $A_1(w_1|\pi)=0$ 表示。同样地，$A_2(w_2|\pi,w_1)=1$表示收益 π 的公司在第一周期拒绝 w_1 之后在第二周期接受开价 w_2，而 $A_2(w_2|\pi,w_1)=0$ 则表示在上述环境下公司在第二周期拒绝 w_2。于是公司的策略由一对函数组成：$[A_1(w_1|\pi),A_2(w_2|\pi,w_1)]$。这里，讨价还价模型的一个特点是公司在贯穿博弈的整个过程中拥有完全信息，因此它的信念是平凡的，公司的信息是私人信息，不为工会所知道，不拥有完全信息的工会在该博弈中首先行动。

根据定义，$\{[w_1,w_2(w_1)],[A_1(w_1|\pi),A_2(w_2|\pi,w_1)],[\mu_1(\pi),\mu_2(\pi|w_1)]\}$必须满足 R_2、R_3 与 R_4 才是完美 Bayes 均衡。我们将证明工会——公司讨价还价博弈存在唯一完美 Bayes 均

衡。所采用的办法是后退归纳法，这是因为公司拥有完美信息且行动在后，他在不知道工会究竟提出怎样的决策的情况下就可以作出自己的最优策略，也就是说它可以最终决定，什么样的开价可以接受，什么样的开价必须拒绝。一般地，我们可以将这种情况归纳如下：如果不完全地告知信息的局中人在拥有完全信息的局中人之前行动，并决不颠倒行动顺序，那么博弈可以试用后退归纳法求解。

我们应用 R_2 于公司的第二周期决策 $A_2(w_2|\pi,w_1)$，当且仅当 $\pi\geqslant w_2$ 时公司的最优决策是接受 w_2（这里像以往一样，我们总是认为等号成立意味着“接受”），w_1 与这个最优决策的充要条件不相干（w_1 只与是否可能进入第二周期讨价还价有关系）。在给定公司在第二周期的最优决策下，再应用 R_2 于工会在该周期有关工资开价的选择。在给定信念 $\mu_2(\pi|w_1)$ 与 $A_2(w_2|\pi,w_1)$ 下，w_2 应当极大化工会的期望效用。在我们的讨论中，最微妙的部分在于信念 $\mu_2(\pi|w_1)$ 的确定。

暂时考虑一个周期的讨价还价模型，因为这样的结果很容易地且合理地被我们看作为两个周期讨价还价中第二周期的解。假设工会相信公司的盈利服从于$[0,\pi_1]$上的均匀分布，这里暂时认为 π_1 是某个任意正数。如果工会开价 w，那么公司的最优反应是显然的：当且仅当 $\pi\geqslant w$ 时接受 w。于是，工会面临的问题可以陈述为求解 w：

$$\max_{w} w\cdot \text{Prob}\{公司接受 w\}+0\cdot \text{Prob}\{公司拒绝 w\} \tag{14.16}$$

由假设 $\pi\sim R[0,\pi_1]$，因此

$$\text{Prob}\{公司接受 w\}=\text{Prob}\{\pi\geqslant w\}=\frac{\pi_1-w}{\pi_1} \tag{14.17}$$

(14.16)式成为

$$\max_{w} w(\pi_1 - w)/\pi_1 \qquad (14.18)$$

由一阶条件易得 $w^*(\pi_1)=\pi_1/2$(当然是在 $w\in[0,\pi_1]$条件下求解(14.18))。

现在重新回到二周期的讨价还价模型，对于任意的一对 w_1 与 w_2，如果工会在第一周期提出工资要求 w_1，而公司预计工会在第二周期的工资开价为 w_2，那么，公司倘若接受 w_1，它的可能盈利为 $\pi-w_1$，而若公司拒绝 w_1，并接受 w_2，那么，它从中获得的可能盈利为 $\delta(\pi-w_2)$，但是如果它对两次开价 w_1 与 w_2 都拒绝的话，它只能获得 0。因此，如果

$$\pi-w_1>\delta(\pi-w_2)$$

或

$$\pi>\frac{w_1-\delta w_2}{1-\delta}\equiv\pi^*(w_1,w_2) \qquad (14.19)$$

公司愿意接受 w_1 而不喜欢等到第二次周期去接受 w_2，而当 $\pi-w_1>0$ 时，公司乐意接受 w_1 而不愿意拒绝两个开价。综上所述，对于任意的两个值 w_1 与 w_2，对于那些效益为$\pi>\max\{\pi^*(w_1,w_2),w_1\}$类型的公司，他们愿意接受 w_1；反之，若 $\pi<\max\{\pi^*(w_1,w_2),w_1\}$ 的话，公司将拒绝 w_1。从实际的情况来陈述，那些效益充分高的公司将在第一周期就接受 w_1，而其他类型的公司则拒绝 w_1。因为 R_2 要求在给定局中人以后的策略时公司作出最优反应，因此，对任意的 w_1 值，如果工会在第二周期的开价为 $w_2(w_1)$的话，公司的最优策略 $A_1(w_1|\pi)$应为

$$A_1(w_1|\pi)\begin{cases}1 & 若\ \pi>\max\{\pi^*(w_1,w_2(w_1)),w_1\}\\0 & 若\ \pi<\max\{\pi^*(w_1,w_2(w_1)),w_1\}\end{cases} \qquad (14.20)$$

现在我们可以得出 $\mu_2(\pi|w_1)$——如果第一周期的 w_1 被拒绝之后，在第二周期到达的信息集上工会的信念。根据完美 Bayes 均衡的要求，信息集上的信念应由 Bayes 法则和局中人可能的均衡策略来确定(R_4)，在第二周期的信息集上，由(14.20)式，工会的信

念必定是对 $A_1(w_1|\pi)=0$ 的那些 π，即 $\pi<\max\{\pi^*(w_1,w_2(w_1)),w_1\}$，认为这些类型均匀地分布着，因此 $\mu_2(\pi|w_1)$ 为 π 服从 $[0,\max\{\pi^*(w_1,w_2(w_1)),w_1\}]$上的均匀分布。给定 $\mu_2(\pi|w_1)$，由前面一周期讨价还价模型的结果知，工会在第二周期的最优开价工资为 $w^*(\pi_1)=\pi_1/2$，这里 $\pi_1=\max\{\pi^*(w_1,w_2(w_1)),w_1\}$。于是我们得到了隐函数方程：

$$\pi_1=\max\{\pi^*(w_1,\pi_1/2),w_1\} \tag{14.21}$$

求解类似(14.21)式的隐函数方程有时是比较困难的，尤其是记号 max 的存在。为去除该记号，先考虑 $w_1\geqslant\pi^*(w_1,\pi_1/2)$的情况，此时必有 $\pi_1=w_1$，但是却有 $\pi^*(w_1,\pi_1/2)=\pi^*(w_1,w_1/2)=\frac{1}{1-\delta}(w_1-\delta\cdot\frac{w_1}{2})=\frac{1-\frac{\delta}{2}}{1-\delta}w_1>w_1$，这与假设 $w_1\geqslant\pi^*(w_1,\pi_1/2)$是矛盾的。于是，在（14.21）式中，我们仅需考虑情况 $w_1<\pi^*(w_1,\pi_1/2)$。此时应有

$$\pi_1=\pi^*(w_1,\pi_1/2)=\frac{1}{1-\delta}(w_1-\delta\cdot\frac{\pi_1}{2}) \tag{14.22}$$

即

$$\pi_1(w_1)=\frac{2w_1}{2-\delta} \tag{14.23}$$

$$w_2(w_1)=\frac{\pi_1}{2}=\frac{w_1}{2-\delta} \tag{14.24}$$

目前我们已经将博弈简化为工会的单周期最优问题：给定了工会在第一周期的工资开价 w_1，我们已经确定了公司在第一周期的最优反应、工会进入第二周期的信念、工会在第二周期的最优开价，以及公司在第二周期的最优反应。于是，应当选择工会的第一周期工资开价为(求解(14.25)式)

$$\max_{w_1}[w_1 \cdot \text{Prob}\{\text{公司接受 } w_1\} + \delta w_2(w_1) \cdot \text{Prob}\{\text{公司拒绝 } w_1 \text{ 但接受 } w_2\} + \delta \cdot 0 \cdot \text{Prob}\{\text{公司拒绝 } w_1 \text{ 与 } w_2 \text{ 这两个开价}\}] \tag{14.25}$$

我们应充分注意到 Prob{公司接受 w_1}不是简单的 π 超过 w_1 的概率;而是 π 超过 $\pi_1(w_1)$的概率

$$\text{Prob}\{\text{公司接受 } w_1\} = \frac{\pi_H - \pi_1(w_1)}{\pi_H} \tag{14.26}$$

注意(14.26)式中我们用了 π_H 而不是一般的 π,其理由是在前面我们已经证明了"只有效益高的公司才可能接受工会的第一周期 w_1 而其他公司都不会接受 w_1"这样一个事实。为求(14.25)式的解,现在的关键在于计算第二项的概率。公司拒绝 w_1 但接受 w_2 当且仅当 $\pi \in (w_2, \pi_1(w_1)) = (\frac{w_1}{2-\delta}, \frac{2w_1}{2-\delta})$时。因此

$$\begin{aligned}
&\text{Prob}\{\text{公司拒绝 } w_1 \text{ 但接受 } w_2\} \\
&= P\{\text{公司在第二周期接受 } w_2 \mid \text{公司拒绝 } w_1\} \cdot P\{\text{公司拒绝 } w_1\} \\
&= \frac{\pi_1}{\pi_H} \cdot \frac{\pi_1 - w_2}{\pi_1} \\
&= \left(\frac{2w_1}{2-\delta} - \frac{w_1}{2-\delta}\right) \Big/ \pi_H \\
&= \frac{w_1}{\pi_H(2-\delta)}
\end{aligned} \tag{14.27}$$

(14.25)式中,中括号里的内容为

$$w_1\left(1 - \frac{2}{\pi_H(2-\delta)} w_1\right) + \delta \cdot \frac{w_1^2}{(2-\delta)^2} \cdot \frac{1}{\pi_H} \tag{14.28}$$

对 w_1 求导并使之为 0,得

$$1 - w_1\left(\frac{4}{2-\delta} - \frac{2\delta}{(2-\delta)^2}\right)\frac{1}{\pi_H}$$

$$=1-\frac{1}{\pi_H}\cdot\frac{2(4-3\delta)}{(2-\delta)^2}\cdot w_1$$
$$=0 \tag{14.29}$$

即

$$w_1^*=\frac{(2-\delta)^2}{2(4-3\delta)}\pi_H \tag{14.30}$$

这就是(14.13)式，由此可得 $\pi_1(w_1^*)$ 与 $w_2(w_1^*)$，其结果如(14.14)式及(14.15)式。

上面所讨论的讨价还价模型是信息不完全的一方先行动(提出开价)，可以想象，如果由拥有完全私人信息的一方先行动的话，其完美 Bayes 均衡的求解有可能发生变化，下面我们以一个简单的"庭外解决"为例展开讨论。

假想一个医生为一个病人切除阑尾炎，在手术过程中发生病人突然死亡事件。病人家属指责医生玩忽职守从而治疗失当酿成悲剧。倘为真，医生将被法庭判决支付 100 万元的高额赔偿金。在上法庭之前他们可以谋求庭外解决，通过谈判达成协议，由医生支付一定数量的钱款给病人家属从而使病人家属撤诉。不管该赔款是庭外解决的结果还是法庭判决的结果，病人家属的律师将分享 25%赔款作为律师费用，当然如果病人家属输掉官司，律师也跟着一无所得。

有一个双方都明白的共同知识：在手术室里当病人死亡时只有医生了解所发生的事。因此，如果法庭调查并需要通过实验来确定是否医疗事故时，医生其实比病人家属更清楚审判结果将是怎样。当然，一旦案件进入审判，医生也必须为此雇请一位律师从而花费 0.1(百万元)。

在庭外讨价还价博弈中，假定由医生(甲)先行动，提出赔偿额 S，病人家属(乙)作最后决定：接受 S 或者拒绝 S 并请求法庭解决。由于乙有不完全信息，仍应用 Harsanyi 转换，让"自然"赋予甲

0.5 概率为有罪，0.5 概率为无辜。我们将一个周期的讨价还价模型的过程用博弈树表示如图 14.8（a 表示接受，R 表示拒绝）所示。

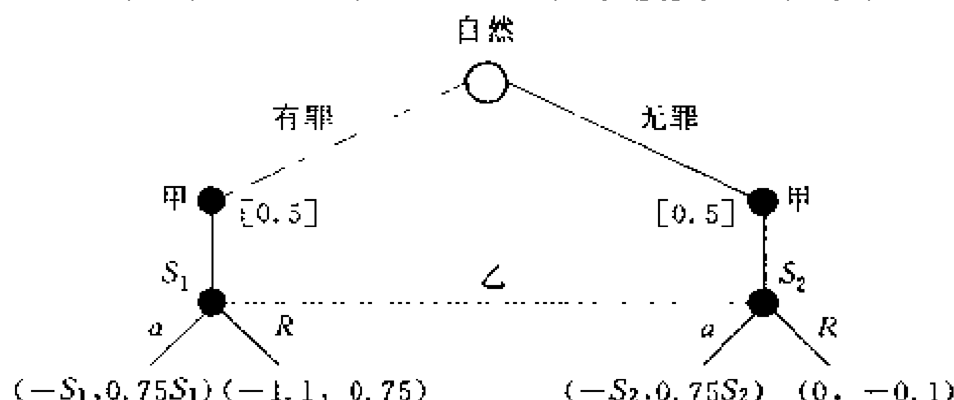


图14.8

图 14.8 中盈利向量(0，−0.1)表示，倘若甲无罪，庭外解决不成而付之于法庭时，我们假设乙必须为甲支付 0.1 的律师费用。

医生甲有两个决策结位于分开的信息集，他的一个纯策略包含两个开价(s_1, s_2)，分别对应于当他为有罪或无罪时作出的允诺。例如(1,0)就是一个可能的策略，当甲知道自己有罪，他在庭外允诺赔偿 100 万元以满足乙方的要求，当甲感到自己无罪，那么他一个子儿也不给乙方。对于病人家属来说，他具有一个信息集，对应于甲可能作出的每一个开价 s。因此，他的纯策略应是 s 的函数，且只取两个值：1——接受；0——拒绝。例如，$A(S)$可能取

$$A(s)=\begin{cases}1 & \text{若 } s\geqslant 1 \text{ 或 } s\leqslant 0.1\\ 0 & \text{当 } 0.1<s<1\end{cases} \tag{14.31}$$

当然，(14.31)式只不过是乙方可能的一个策略，拒绝域是否一定是 $s\in(0.1,1)$，有待于进一步讨论。事实上，我们需要寻求完美 Bayes 均衡解，看看乙方的 $A(S)$究竟如何。为此，我们必须探讨在给定甲方提出的 S 条件下，乙方认为甲有罪的信念。显然这个信念是 S 的函数，记作 $\pi(S)$。例如，可以平凡地认为 $\pi(S)=0.5$ 对所有开价 S 成立。也可以认为，当 $S\geqslant 1$ 时，$\pi(S)=1$，当 $S<0.1$ 时，$\pi(S)=0$。对于其余所有的 S 有 $\pi(S)=0.5$。我们也需要确定构成

完美 Bayes 均衡中的病人家属乙的信念。

为寻求完美 Bayes 均衡解，在这个博弈模型中，显然后退归纳法不能奏效，因为在诸终点结上面一步的信息集集合具有连续统势。

如果后退归纳法不能使用的话，不妨尝试从通过寻求每一个局中人的劣策略并剔除所有劣策略开始着手。

在这个博弈模型中，一个基本常识是，如果法庭判乙方胜诉，乙方至多获得 100 万元赔偿。因此开庭前甲方提出的任何超过(包含)1(百万元)的赔偿金都应当被乙方接受。这个策略总是(弱)优于任何拒绝超过赔偿金 1 的策略。之所以称拒绝 1 以上的策略是弱劣的，因为在给定了医生甲的某种开价(例如 $S=0.3$)，“接受 1 以上”并不比“拒绝 1 以上”有更多获益。现在我们剔除乙方的弱劣策略，讨论仅限于

$$A(s)=\begin{cases}? & 若\ s<1 \\ 1 & 若\ s\geqslant 1\end{cases} \tag{14.32}$$

因为乙接受 1(包括 1)以上的任何开价，那么对于甲来说，“开价超过 1”的策略无疑地劣于“开价恰为 1”的策略。然而，另外一个基本常识又告诉我们，如果甲的确无罪，他将不愿付出任何赔偿金，即他不会因此面失去什么(当然，所花费的时间与精力不在本模型考虑之内)。用劣策略的语言来说，如果甲无罪，“赔偿任何钱款给乙”的策略应当弱劣于“甲拒绝庭外解决”的策略。于是，我们仅需要将讨论限于甲的纯策略形式：$(S_1,0)$，其中 $S_1\leqslant 1$。

为求完美 Bayes 均衡解，当每次我们知道拥有信息的局中人某些策略时，不妨看看这对于无信息局中人的信念意味着什么。

例如，如果医生甲曾经提出过非零的赔偿金用于庭外解决，那么病人家属乙显然将推断甲在一定程度上是有罪的。因此它关于甲的信念将由此修正为

$$\pi(s)=\begin{cases}\theta & 若\ s=0\\ 1 & 若\ s>0\end{cases} \tag{14.33}$$

其中 θ 待确定。

在求解完美 Bayes 均衡的过程中，每次修正无信息局中人信念时，总是剔除那些目前为劣的策略。

根据这个原则，病人家属乙知道，每当甲提供赔偿金时，他可以通过法庭赢得 100 万元。因此乙应当拒绝 0 至 100 万元之间的任何开价。因而对于乙来说，可以有两个非劣纯策略。

较懦弱一些的纯策略

$$A(s)=\begin{cases}1 & 若\ s=0\ 或\ s\geq 1\\ 0 & 若\ 0<s<1\end{cases} \tag{14.34}$$

较强硬一些的纯策略

$$A(s)=\begin{cases}1 & 若\ s\geq 1\\ 0 & 若\ s<1\end{cases} \tag{14.35}$$

这两种纯策略的区别在于 $s=0$ 时的决策。因为当 $s>0$ 时，根据纠正了的信念，乙相信甲是有罪的，由剔除劣策略原则，$0<s<1$ 对于乙来说是不可取的，而当 $s=0$ 时，懦弱的乙认为甲可能无罪，强硬的乙仍认为甲有罪的可能性较大，这些取决于待定的 θ。

在我们为一个局中人剔除劣策略时，还应当检查一下，看看这一举措是否可能为另一个局中人制造出新的劣策略。

倘若医生甲知道自己是有罪的，那么他明白上法庭将花去他 1.1(百万元)，因此任何从 0 至 1 之间的提议都将被乙所拒绝，因为这些策略显然劣于他出价 1 的策略，出价 1 为乙所接受从而使甲少出了 0.1。现在，对于甲来说，我们有两个非劣的纯策略：$B=(0,0)$，$F=(1,0)$。前者其实是一个共有策略，而后者是一个分离策略。

如果我们已经剔除了所有的(累次)劣策略，由此将每一个局中人的策略缩减到少数几个，那么计算这个"简化"了的博弈的策略型并求出它的 Bayes 均衡。如果这样的 Bayes 均衡只有一个，它

就是我们所要求的完美 Bayes 均衡。

在我们的例子中，盈利矩阵如图 14.9。

		甲(有罪)		甲(无罪)	
		B	F	B	F
乙	懦弱	0,0	0.75,−1	0,0	0,0
	强硬	0.75,−1.1	0.75,−1	−0.1,0	−0.1,0

图 14.9

考虑到医生甲有罪与无罪的先验概率各为 0.5，因此我们得到“简化”博弈的策略型如图 14.10 所示。

		甲	
		B	F
乙	懦弱	0,(0,0)	0.375,(−1,0)
	强硬	0.325,(−1.1,0)	0.325,(−1,0)

图 14.10

图 14.10 中每格的第一个元素表示乙取相应策略时的期望盈利，第二个(向量)元素表示甲取相应策略时的盈利向量，其第一个元素表示甲有罪时的盈利，第二个元素表示甲无罪时的盈利，读者不妨一一验证之。由图 14.10 不难看出，它不存在纯策略的完美 Bayes 均衡。现在令医生甲以概率 p 取 B，以概率$(1-p)$取 F，而乙以概率 q 取 W(懦弱)，以概率$(1-q)$取 T(强硬)。我们在不完全信息静态博弈中已经提到，这样的混合策略可以与局中人的信念相联系。

我们利用 Bayes 法则来计算乙的信念，当甲拒绝庭外解决(甲的开价为 0)时，乙认为甲有罪的信念为

$$\pi(0)=P(\text{甲有罪}\mid s=0)$$

$$=\frac{P(s=0\mid\text{甲有罪})\cdot P(\text{甲有罪})}{P(s=0\mid\text{甲有罪})\cdot P(\text{甲有罪})+P(s=0\mid\text{甲无罪})P(\text{甲无罪})}$$

$$=\frac{p\cdot 0.5}{p\cdot 0.5+1\times 0.5}=\frac{p}{1+p} \tag{14.36}$$

这里，P(甲有罪)$=P$(甲无罪)$=0.5$ 为先验概率。显然，当甲知道自己无罪时他的开价 $s=0$ 的概率为 1，而当甲有罪时却开价 $s=0$ 这件事相当于甲取策略 B，故条件概率为 p。(14.35)式实际上确定了(14.32)式中的 θ。

混合策略的求解，即求 p 使得乙方在选择 W 或选择 T 之间取无所谓态度，即

$$p\cdot 0+(1-p)\cdot 0.375=0.325p+0.325(1-p)=0.325 \tag{14.37}$$

得

$$p=\frac{2}{15}$$

这意味着，在缩减了的"简化"博弈中的医生甲的均衡策略应为：如果甲有罪，他以 $p=\frac{2}{15}$ 概率开价 0，而以概率 $\frac{13}{15}$ 开价为 1；如果甲无罪，那么他总是开价为 0。

现在我们需要确定乙的均衡策略，也就是需要确定 q。同样理由，q 的合理确定应使甲在策略 B 与 F 之间的选择方面感到真正的无所谓，由于甲是否有罪的先验概率各为 0.5，因此他取 B 与 F 时分别可能得到的盈利可以用如图 14.11 所示的矩阵来概括。

甲的期望盈利

		B	F
乙	W	0	-0.5
	T	-0.55	-0.5

图 14.11

于是，q 应当满足

$$q\cdot 0-0.55(1-q)=-0.5q-0.5(1-q)=-0.5 \tag{14.38}$$

得

$$q=\frac{1}{11}$$

“简化”博弈的唯一均衡是我们所要求的完美 Bayes 均衡：

医生甲：如果有罪，策略为($B(s=0)$：概率为$\frac{2}{15}$，$F(s=1)$：概率为$\frac{13}{15}$)。

如果无罪，策略为 $s=0$。

病人家属乙方：

$$A(s)=\begin{cases}以概率\frac{1}{11}为 1，若 s=0\\0 \quad 若 0<s<1\\1 \quad 若 s\geqslant 1\end{cases}$$

乙方的信念：

$$\pi(s)=\begin{cases}\frac{2}{17} \quad 若 s=0\\1 \quad 若 s>0\end{cases}$$

在这个例子中，我们以赔款额的大小直接作为双方的可能盈利，如果从其他角度重新定义合理的效用函数的话，那么完美 Bayes 均衡将可能有所变化。

第十五章　完美 Bayes 均衡的精炼

我们已经树立了有关对均衡进行精炼的概念，精炼的思想很容易推广到完美 Bayes 均衡。也就是在进一步条件或限制下，试图得到更为合理的博弈预测结果。在不完全信息动态博弈中，我们考虑了在非均衡路径有关信念的进一步要求，从而达到对完美 Bayes 均衡的精炼。

§15.1　剔除严劣策略

根据完美 Bayes 均衡的四个要求，任何一个局中人 i 在任何信息集的开头不可以取严劣策略，因此对于局中人 $j(\neq i)$ 来说，如果去相信局中人 i 将采取严劣策略显然是不合情理的。我们用一个简单的例子来使这种想法具体化。

考虑图 15.1 所示博弈：

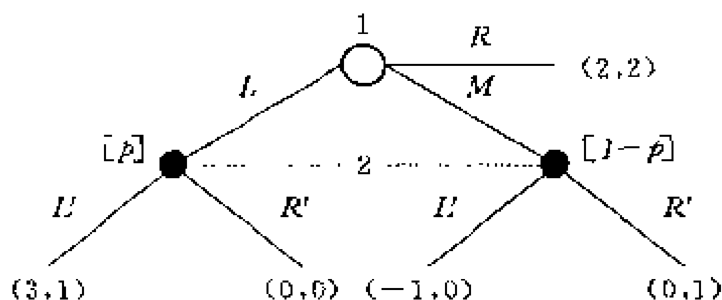


图15.1

在不考虑局中人 2 在自己行动的信息集上的信念时，图 15.1

可以写成如图 15.2 的策略型表示。

		局中人 2 L'	R'
	L	3,1	0,0
局中人 1	M	1,0	0,1
	R	2,2	2,2

图 15.2

该正则型博弈有两个纯策略 Nash 均衡：(L,L')与(R,R')。图 15.1 展开型博弈除了本身外没有其他的子博弈。于是(L,L')与(R,R')当然是子博弈完美均衡。读者可以注意到，在均衡(L,L')中，局中人 2 的信息集位于均衡路径上，因此根据完美 Bayes 均衡要求 R_3，$p=1$。于是$(L,L',p=1)$构成了博弈的纯策略完美 Bayes 均衡。除此之外，事实上还存在另一个纯策略完美 Bayes 均衡：$(R,R',p\leqslant 1/2)$。因为在均衡(R,R')中，局中人 2 的信息集不在均衡路径上，R_4 对 p 没有什么限制。于是我们对于局中人 2 的信念 p 的要求，仅仅是使得 R'成为最优行动，也就是说，$p\leqslant 1/2$（即，$1-p>1/2$）。这个例子的一个关键特点是：M 其实是局中人 1 的严劣策略。图 15.2 明确地告诉我们，M 严劣于 R。因此，"局中人 2 相信局中人 1 可能取行动 M"显然是不合理的。用数学语言正式地叙述，信念 $1-p>0$ 是不合理的。于是，从纯策略完美 Bayes 均衡中剔除不合理的$(R,R',p\leqslant 1/2)$。$(L,L',p=1)$是唯一满足我们在本节所提出（精炼）要求的完美 Bayes 均衡。再一次叙述这个要求：局中人 j 不会认为局中人 i 会采取严劣策略。

现在我们考虑这样的情况，将局中人 1 在(L,L')之后结局的盈利由 3 改为 3/2，由图 15.2，此时，非但 M，而且 L 都是局中人 1 的严劣策略。与前面一样地进行讨论，局中人 2 不会相信局中人 1 会取严劣策略 L。或者说，$p>0$ 是不合理的，p 必须为 0。这样就与前面我们所得到的 $p=1$ 的结果矛盾。在这种情况下，本节所提出

的要求对局中人 2 在非均衡路径上的信念将没有什么限制。因为局中人 2 相信，局中人 1 既不会取 L，也不会取 M。

在图 15.1 这个博弈中，我们发现 M 不仅仅是在（局中人 1 的）信息集的起始时为严劣的，而且它在整个博弈中是局中人 1 的严劣策略。注意在这里我们用了两种讲法：严劣和在信息集的起始为严劣。为了正确且直观地理解这两种讲法之间的区别，不妨假设在局中人 1 行动之前，局中人 2 具有行动。从初始结出发，局中人 2 要么结束博弈，要么让局中人 1 如图 15.1 那样去选择行动。我们将这个假想博弈的博弈树表示如图 15.3。

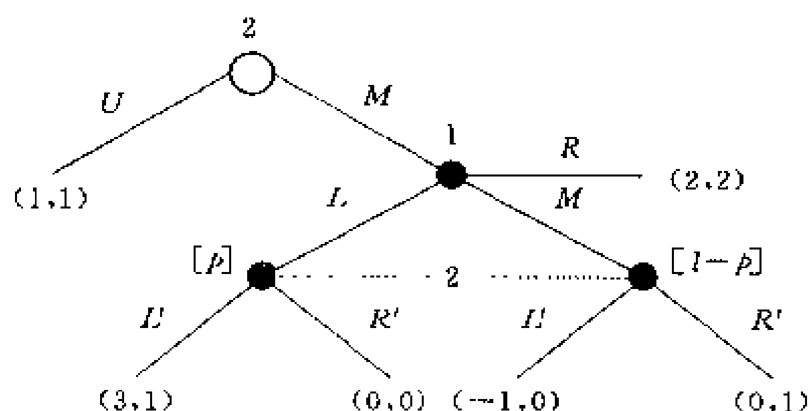


图15.3

在图 15.3 的展开型博弈中，在局中人 1 的信息集的起始，M 仍然是局中人 1 的严劣策略，但就整个博弈来说，M 不是严劣的。因为一旦局中人 2 在初始结采取行动 U 从而结束博弈，此时，对局中人 1 来说，三个策略：L、M、R 全产生相同的盈利。

局中人 i 的一个策略 s_i'，称为是严劣的，则一定存在另一个策略 s_i，使得两者对于与其他局中人的策略的每一种可能的组合，局中人 i 由取 s_i 所得的盈利总是大于由取 s_i' 所得的盈利。在图 15.1 中，M 是局中人 1 的严劣策略，因为无论局中人 2 取什么策略，局中人 1 取 M 的盈利总是小于他取 R 时所得的盈利。但在图 15.3 中，M 不是局中人 1 的严劣策略，因为 L、M、R 与局中人 2 的初始行动 U 的组合产生同样的盈利。

那么在局中人 i 的信息集的起始(或开头)为严劣策略又是怎么回事呢?首先的前提是考虑 i 具有行动的信息集,一个策略 s_i' 称为在该信息集开始时为严劣的,如果存在另一个策略 s_i',使得对于局中人 i 在给定的信息集所拥有的每一个信念(这一点很重要,因为它涉及在这个信念下局中人 i 的最优策略),以及对于与其他局中人的后续策略的每一种可能的组合,局中人 i 从在给定信息集上取 s_i 所确定的行动和由取 s_i 所确定的后续策略所得到的期望盈利必定严格地大于他由取 s_i',所确定的行动和相应的后续策略所得到的期望盈利。请注意上述说法的一个基本事实,所谓在信息集起始时的严劣策略这个概念只考虑在该信息集以后发生的情况。一个符合实际且合理的想法是,我们将坚持这样的要求:局中人 j 不应当相信局中人 i 可能取一个在任何信息集起始时为严劣的策略。这个要求以 R_5 正式地表述如下:

R_5 如果可能,每一个局中人的非均衡路径信念在那些仅当另一个局中人取前面一个信息集起始时为严劣的策略而达到的结上置零概率。

在要求 R_5 中,注意到限定条件"如果可能"的含义。举例说明,在图 15.1 的博弈中,倘若像前面所述,将(L,L')的后续终点结上局中人 1 的盈利 3 改为 3/2,那么,在局中人 1 的纯策略空间中,R 优于 L 与 M。在这种情况,根据原先完美 Bayes 均衡的 R_1,要求局中人 2 在自己的信息集上有一个信念,但是这个信念不可能在 M 与 L 的后续结上都置零概率。这时,其实 R_5 将不适用。也就是说,R_5 不限制局中人 2 在非均衡路径上的信念。

再考虑图 15.4 所示的博弈。

这是一个信号博弈,如第十三章所述,发送者的策略(m',m'')意指类型 t_1 选取信号 m',类型 t_2 选取信号 m''。接收者的策略(a',a'')意指接收者在收到信号 L 之后选取行动 a',收到信号 R 之后选取行动 a''。

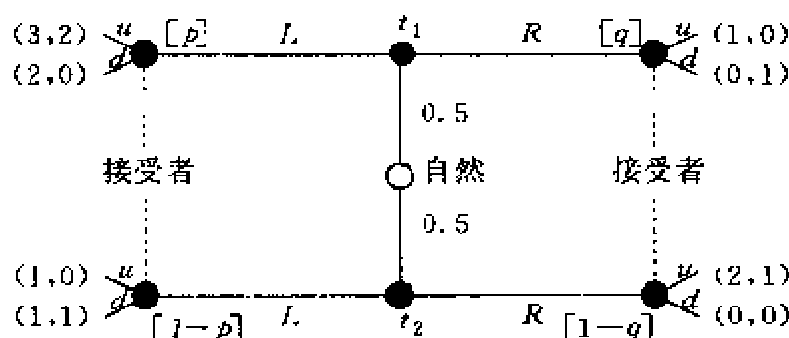


图15.4

首先验证[(L,L),(u,d),$p=1/2$,q]对于任何$q\geq 1/2$均构成一个共有完美 Bayes 均衡(当然是在要求R_1—R_4下)。假如存在一个均衡,其中发送者的策略是(L,L),那么接收者对应于L的信息集当然在均衡路径上,注意到确定发送者类型的先验分布为:类型t_1具有概率 1/2,类型t_2具有概率 1/2。因此由 Bayes 法则,$p=1/2$。于是在获得信号L之后,局中人 2 取u的期望盈利为$1/2\times 2+1/2\times 0=1$,而取$d$的期望盈利为$1/2\times 0+1/2\times 1=1/2$,$u$是信号$L$之后的最优选择行动。此时,类型$t_1$将获益 3,类型$t_2$将获益 1。现在我们需要确定发送者的两个类型$t_1$与$t_2$是否都愿意选用"共有"信号$L$。对于类型$t_1$来说,无论接收者的反应是$u$还是$d$,他取$L$分别获盈利 3 或 2,总是大于他取$R$时的分别获益 1 或 0。因此,类型$t_1$必定愿意取$L$。再来看类型$t_2$,如果他取$L$,已知接收者的最优反应是$u$,$t_2$得 1。如果$t_2$取$R$,接收者的反应若为$u$的话,$t_2$得 2,若接收者反应为$d$的话,$t_2$获利为 0。为使$t_2$也不偏离策略$L$,那么接收者对于$R$的反应应当是$d$。因此$(L,L)$若为均衡中发送者的策略,那么接收者的策略一定是$(u,d)$。现在我们考虑接收者在对应于$R$信息集上的信念,以及在给定这个信念$(q,1-q)$下取$d$的最优性。显然当$q\geq 1/2$时接收者对$R$的最优反应为$d$,因此当$q\geq 1/2$时,[$(L,L)$,$(u,d)$,$p=0.5$]构成了该信号博弈的共有完美 Bayes 均衡。但是,我们已经分析到,在这个博弈中,t_1取R是毫无意义的。发送者的策略(R,L)和(R,R)——这

表明，t_1 总取 R—— 在对应于 t_1 类型发送者的信息集起始时为严劣的。例如，策略(L,R)使得 t_1 获得的盈利至少为 2，但是(R,L)或(R,R)使 t_1 获得盈利至多为 1。所以，接收者在看到信号 R 之后的信息集可以在相应于 t_2 的那个结上通过发送者的策略(L,R)(它至少优于(R,L)和(R,R))而达到。但是在相应于 t_1 的那个结上却不能达到。因此，R_5 就要求 $q=0$。尽管$[(L,L),(u,d),p=0.5]$当 $q\geqslant 1/2$时为完美 Bayes 均衡，但是，它却不能满足 R_5 的要求。换句话说，完美 Bayes 均衡$[(L,L),(u,d),p=0.5,q\geqslant 1/2]$由于不符合 R_5 而从均衡解中被精炼掉。

那么该信号博弈是否存在满足 R_5 要求的完美 Bayes 均衡呢？分离完美 Bayes 均衡$[(L,R),(u,u),p=1,q=0]$就由于不存在非均衡路径的信息集而"平凡地"满足了 R_5，从而免去了被 R_5 精炼掉的下场。我们顺便证明这个策略是分离完美 Bayes 均衡：如果发送者取分离策略(L,R)，那么接收者在均衡路径上的信念必然分别为 $p=1$ 和 $q=0$，由图 15.4 可知，接收者的最优反应均是 u，并分别获得盈利 2 与 1。如前面所分析的那样，类型 t_1 不会偏离 L。而类型 t_2 也不应当偏离 R，因为 t_2 若取 L，由于接收者的策略为(u,u)，此时 t_2 只获盈利 1，比起他取 R 时的盈利 2 少了 1。

但是，从上述信号博弈的结果不要产生这样的误解：在信号博弈中，共有完美 Bayes 均衡由于在非均衡路径上的信念受到 R_5 的限制而必定成为被精炼的对象。事实并非如此。如果将图 15.4 信号博弈中当类型 t_2 发送者发送 R 之后，接收者关于行动 u 与 d 的相应盈利颠倒一下，即，接收者此时取 u 获利 0 而取 d 获利 1。利用前面同样的讨论，易知$[(L,L),(u,d),p=0.5,q$ 取$[0,1]$中的任意数$]$构成了信号博弈的共有完美 Bayes 均衡。因为无论 q 取任何数，只要 $q\in[0,1]$，在接收到信号 R 之后，接收者的最优反应一定是 d。特别地，我们取 $q=0$，完美 Bayes 均衡$[(L,L),(u,d),p=0.5,q=0]$满足 R_5 要求，这是一个非平凡地满足 R_5 要求的共有完

美 Bayes 均衡的例子。

本节中提出 R_5 要求以对均衡进行进一步精炼，我们所举的例子是信号博弈。鉴于信号博弈本身的重要性和广泛的应用性，我们不妨叙述将 R_5 用于信号博弈中的完美 Bayes 均衡的等价形式。

首先叙述严劣信号概念：

定义 15.1 在信号博弈中，类型 t_i 的来自信号空间 M 的信号 m_j 称为劣的，如果存在 M 中的另一个信号 $m_{j'}$，使得 t_i 由取 $m_{j'}$，所得的最低可能盈利大于 t_i 取 m_j 所得的最高可能盈利：

$$\max_{a_k \in A} U_s(t_i, m_{j'}, a_k) > \max_{a_k \in A} U_s(t_i, m_j, a_k) \qquad (15.1)$$

在图 15.4 中，类型 t_1 的 R 是劣的，因为他取 R 将最高获得盈利 1，而他取 L 的最低可能盈利为 2。读者不难验证，对于类型 t_2 来说，L 和 R 都不是劣的。

SR_5 在信号博弈中，如果 m_j 之后的信息集位于非均衡路径，且 m_j 是类型 t_i 的劣策略，那么(如果可能的话)，接收者的信念 $\mu(t_i|m_j)$ 应当对类型 t_i 置零概率(假如 m_j 对类型空间 T 中的所有其他类型并不是劣的话，这种情况是可能的)。

$[(L,L),(u,d),p=0.5,q]$ 对一切 $q\in[0,1]$ 都是略修改之后的图 15.4(t_2 发送 R 后，接收者相应于 u 于 d 的盈利互换位置)信号博弈的共有完美 Bayes 均衡，对 t_1 来说，R 是劣的，但 R 对 t_2 并非劣的，因此在 R 之后的信息集上，根据 SR_5，接收者的信念 $\mu(t_1|R)=0=q$。

§15.2 "啤酒与蛋奶火腿蛋糕"信号博弈

R_5 对均衡的精炼作用似乎是无可非议的，然而紧接着会出现两个令人们关心的问题：

(1)是否存在这样的完美 Bayes 均衡，它看上去不怎么合理，

但是它仍然满足 R_5? 一个完美 Bayes 均衡在怎样情况下会成为不合理的呢?

(2)对于均衡的定义,需要再加上些什么样的要求,才可以剔除那些不合理的完美 Bayes 均衡呢?

Cho 和 Kreps 于 1987 年构造了一个例子——"啤酒与蛋奶火腿蛋糕"(Beer and Quiche)信号博弈,从而阐述了不合理的完美 Bayes 均衡可以满足 R_5。

在"啤酒与蛋奶火腿蛋糕"信号博弈中,发送者有两个类型:t_w 表示懦弱(weak),t_s 表示粗暴(surly),"自然"为发送者选取类型的先验分布是:发送者取 t_w 的概率为 0.1,发送者取 t_s 的概率为 0.9。如果接收者相信发送者是懦弱的(其概率超过 1/2),那么他愿意跟懦弱的发送者斗,但是接收者不愿意与粗暴的发送者斗。接收者无法观察到发送者属于何种类型,然而他可以在作出斗与不斗的决策之前观察到发送者以什么东西充作早餐。发送者的早餐只有两种选择:"啤酒"和"蛋奶火腿蛋糕"(不可兼而得之);粗暴类型的发送者喜欢啤酒,懦弱类型的发送者喜欢蛋奶火腿蛋糕。两种类型的发送者都不愿意与接收者发生争斗,因此他们关心是否争斗甚于关心自己所用早餐是什么,但他们所用早餐却成为接收者可观察的信号。这个信号博弈的展开型表示及相应盈利向量如图 15.5 所示。

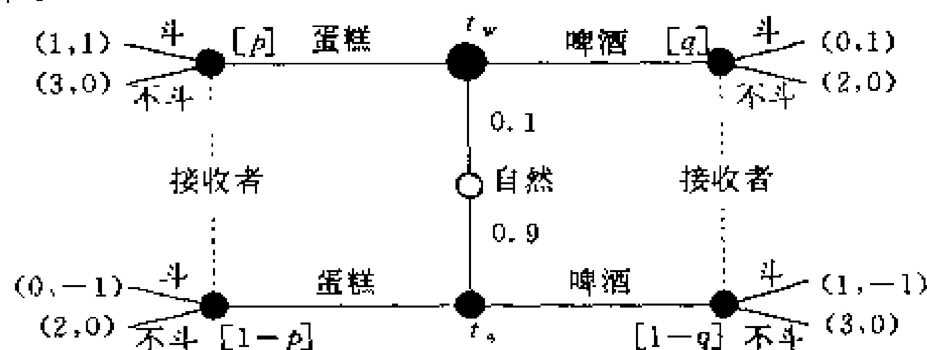


图15.5

在"啤酒与蛋奶火腿蛋糕"信号博弈中,[(蛋糕,蛋糕),(不斗,

斗），$p=0.1,q$]对任意 $q\geqslant 1/2$ 构成共有完美 Bayes 均衡。（注：发送者策略向量的第一个元素相应于类型 t_w 的行动，第二个元素则表示类型 t_s 采取的行动；策略剖面中的第二个向量表示接收者的策略行动，其中第一个元素是接收者观察到信号“蛋糕”之后采取的行动，第二个元素则表示接收者观察到信号“啤酒”之后所采取的行动。）先来验证这个事实：假如存在一个均衡，其中发送者的策略是（蛋糕，蛋糕），那么接收者对应于“蛋糕”的信息集必定在均衡路径上，利用 Bayes 法则与已知的先验分布，可知，$p=0.1$。对于类型 t_w 来说，“蛋糕”优于“啤酒”，因为如果他取“蛋糕”，他的盈利为 1（若接收者“斗”）或者为 3（若接收者“不斗”），而如果他取“啤酒”，相应的盈利分别为 0 或 2。因此类型 t_w 不会偏离该策略剖面。至于类型 t_s 会不会偏离呢？这就需要考虑接收者的反应。如果发送者的信号是“蛋糕”，由于 $p=0.1$，接收者取“斗”的期望盈利为 $1\times 0.1+(-1)\times 0.9=(-0.8)$，少于取“不斗”时的期望盈利 0，因此“不斗”是接收者关于信号“蛋糕”的最优反应。也就是说，t_s 发出信号“蛋糕”的话，他可以指望获得盈利 2。如果发送者在早餐时选择“啤酒”，倘若接收者的反应是“斗”，那么 t_s 将获得盈利 1，比 2 少了 1。因此在给定接收者（不斗，斗）的策略时，类型 t_s 不会偏离。最后我们考虑接收者在非均衡路径上的信念，为使在这个信念下接收者的最优策略是“斗”，显然只要令 $q\geqslant 1/2$ 就足够了。

现在我们指出这个完美 Bayes 均衡是满足 SR_5 的。这是因为对于发送者的任何一种类型，“啤酒”不是劣策略。这个断言对于 t_s 来说并不难理解，因为接收者取“斗”的话，t_s 取“啤酒”获利为 1，而若取“蛋糕”则获利仅为 0。关键在于该断言是否适合于类型 t_w，从图 15.5 可以看出，t_w 不能担保自己若取“蛋糕”其结果一定会更好一些，此时他的最差可能盈利为 1，而他若取“啤酒”其最好可能获利为 2。正因为“啤酒”不是劣的，因此 SR_5 不限制 q 的取值。

但是，从另一方面来看，接收者的非均衡路径信念似乎有点可

疑：那就是，如果接收者意外地（因为在非均衡路径上）观察到"发送者的早餐用的是啤酒"这样一个信号，信念 $q \geq 1/2$ 将使接收者断定"发送者类型为懦弱的可能性至少不会小于粗暴类型的可能性"。其实，如下两条推理无疑是合理的：(a)懦弱类型不可能通过早餐用啤酒而不用蛋糕以对均衡盈利 3 有所改善；(b)粗暴类型可以改善他的均衡盈利 2，如果接收者的信念为 $q<1/2$ 则使粗暴类型有可能得到盈利 3。尽管如此，"$q \geq 1/2$"仍使接收者会作出上述推断。

在(a)与(b)这两种推理下，人们很自然地可能预测"粗暴类型在早餐中选用啤酒"，于是发送者应当这样地去推理：只要我在早餐中选择啤酒，那么接收者会相信我是属于粗暴类型的，因为由(a)，选择啤酒不可能改善懦弱类型者获更多盈利；而由(b)，如果选择啤酒将使接收者相信我是粗暴类型的，一旦接收者相信这一点他因为不愿与我斗而采取不斗策略，从而提高了我的盈利。

倘若确信上述推理，这意味着要求 $q=0$，于是与我们在上面确定的共有完美 Bayes 均衡不相容。

我们可以得出如下结论：

[(蛋糕，蛋糕)，(不斗，斗)，$p=0.1, q \geq 1/2$]是博弈的共有完美 Bayes 均衡，它满足 SR_5，但是它不太合理。

其实，在"啤酒与蛋奶火腿蛋糕"信号博弈中还存在着另一个共有完美 Bayes 均衡，那就是[(啤酒，啤酒)，(斗，不斗)，$q=0.1$，$p \geq 1/2$]。假定存在一个均衡，其中发送者的信号选择是(啤酒，啤酒)，由 Bayes 法则，不难确定在均衡路径上的信念为 $q=0.1$。在这个信念下，接收者对"啤酒"的最优反应是"不斗"，这样他至少盈利为 0，否则他的期望盈利为(−0.8)。此时 t_w 获盈利 2，t_s 获盈利 3。为确定发送者的两个类型是否都愿意选"啤酒"，我们必须确定接收者关于"蛋糕"的最优反应。如果接收者关于"蛋糕"的反应是不斗，那么 t_s 从"蛋糕"中获盈利 2，少于他取"啤酒"的可能盈利 3，

但 t_w 却从"蛋糕"获得 3，多于他取"啤酒"的盈利 2。然而，若接收者关于"蛋糕"的反应是"斗"，那么 t_w 与 t_s 从"蛋糕"中分别获得 1 与 0，当然少于他们取"啤酒"时所得的 2 与 3。因此，如果存在一个均衡，其发送者策略是（啤酒，啤酒）的话，那么接收者对"蛋糕"的反应一定是斗，在非均衡路径的信息集上，要使接收者感到取"斗"是最优的，必应有 $p \geq 1/2$。于是我们证明了[（啤酒，啤酒），（斗，不斗），$q=0.1$，p]当 $p \geq 1/2$ 都构成博弈的完美 Bayes 均衡。这个均衡同样地满足 SR_5，因为"蛋糕"对于两种类型的发送者都不是劣策略：t_w 取"蛋糕"的最高盈利为 2，多于他取"啤酒"的最低盈利 1。但是这个完美 Bayes 均衡中，接收者的非均衡信念 $p \geq 1/2$ 不会令人对它产生任何怀疑。看到发送者早餐时选择蛋奶火腿蛋糕，推断他属于懦弱类型的可能性不小于他属于粗暴类型的可能性。这种推断至少在直观上比较合理。

对那种不太合理的但取满足 SR_5 的完美 Bayes 均衡需要另外强加要求（在非均衡路径上）才有可能剔除，此类讨论推广到整个信号博弈类，这就是我们在下一节将要介绍的由 Cho 和 Kreps 于 1987 年提出的直觉准则（the intuitive criterion）。

§15.3 直觉准则

在引入直觉准则之前，我们先介绍一个概念。

定义 15.2 给定一个信号博弈中的某完美 Bayes 均衡，称来自信号空间 M 的信号 m_j 对于类型空间 T 中类型 t_i 为均衡劣（equilibrium-dominated）的，如果 t_i 的均衡盈利 $U^*(t_i)$，大于 t_i 来自 m_j 的最高可能盈利：

$$U^*(t_i) > \max_{a_k \in A} U_s(t_i, m_j, a_k) \tag{15.2}$$

均衡劣是针对信号博弈中给定的完美 Bayes 均衡中发送者的

信号策略而言的，如果一个信号 m_j 使类型 t_i 发送者可能获得的最大盈利小于 t_i 在完美 Bayes 均衡中获得的盈利，想来这个信号策略对于类型 t_i 不怎么吸引人。上一节中，[(蛋糕，蛋糕)，(不斗，斗)，$p=0.1, q\geqslant 1/2$]中"啤酒"是 t_w 类型发送者的均衡劣信号策略，在[(啤酒，啤酒)，(斗，不斗)，$q=0.1, p\geqslant 1/2$]中，"蛋糕"是 t_s 类型发送者的均衡劣信号策略。一个合理且直观的想法是，接收者不应当相信发送者 t_i 会发出均衡劣的信号，我们将这种想法用 SR_6 正式地表述：

SR_6(直觉准则 Cho & Kreps 1987)：如果信号 m_j 后面信息集在非均衡路径上，并且 m_j 对于类型 t_i 的发送者是均衡劣的，那么(如果可能的话)，接收者在该信息集的信念 $\mu(t_i|m_j)$ 应对类型 t_i 置零概率(假如 m_j 并不是对所有 T 中的类型为均衡劣的话，这种情况是可能的)。

"啤酒与蛋奶火腿蛋糕"揭示了这样一个事实：信号 m_j 对于类型 t_i 来说，不必是劣的但可以是均衡劣的。例如，对于懦弱类型 t_w，"啤酒"不是一个劣策略，但是在给定的[(蛋糕，蛋糕)，(不斗，斗)，$p=0.1, q\geqslant 1/2$]这个完美 Bayes 均衡中，"啤酒"对 t_w 是均衡劣的。按照直觉准则，接收者应当不相信 t_w 会发出"啤酒"信号，因为不管 t_w 如何预测接收者对"啤酒"将会作出何种反应，没有什么可以激励 t_w 在早餐时饮用啤酒。一旦在非均衡路径上接收者将信念集中于 t_s，那么接收者的唯一最优反应就是不与粗暴的发送者斗，从而使 t_s 得到比均衡盈利更多的盈利。

然而，如果 m_j 关于 t_i 是劣的话，那么 m_j 必定关于 t_i 是均衡劣的。为证明这一点，仅需结合(15.1)式与(15.2)式，由下述不等式即可看到：

$$U^*(t_i)\geqslant\max_{a_k\in A}U_s(t_i,m_j',a_k)>\max_{a_k\in A}U_s(t_i,m_j,a_k) \quad (15.3)$$

因此，对一个完美 Bayes 均衡强加要求 SR_6 的话，SR_5 其实就成为

多余的要求。利用 Kohlberg 与 Mertens(1986)的一个结果，Cho 与 Kreps 证明了在第十三章中所定义的信号博弈类中的任何信号博弈都有满足 SR_6 的完美 Bayes 均衡。

在 Spence 劳力市场信号模型中存在妒忌的情况中，由一个分离完美 Bayes 均衡是满足 SR_6 的，在那里，低能力工人选择 $e^*(L)$，而高能力工人选取足够的教育水平 e_i，使得低能力工人不在乎去摹仿高能力工人从而“骗”得高工资。有兴趣的读者可以看 Gibbons 的《Game Theory for Applied Economists》的最后一章。

参考文献

1. Beirman, H. Scott and Fernandez, Louis. 1998. Game Theory with Economic Applications. Addision-Welsley

2. Fudenberg, Drew and Jean Tirole. 1991. Game Theory. MIT Press

3. Gibbons, R. 1992. Game Theory for Applied Economists. Priceton Univ. Press

4. McMillan, John. 1992. Games Strategies and Managers. Oxford Univ. Press

5. 张维迎. 1996. 博弈论与信息经济学. 上海:上海人民出版社,上海三联书店

6. 富兰克林著. 俞建,顾悦译. 1985. 数理经济学方法——线性和非线性规则、不动点理论